

June 1984

NASA-RP-1120 19840026261

# Solar Terrestrial Physics: Present and Future

*A Report Based on the Solar  
Terrestrial Physics Workshop,  
December 1982 to November 1983*

FOR REFERENCE

1984 JAN 11 10 11 AM

UNCLASSIFIED COPY

1984  
NASA HEADQUARTERS  
WASHINGTON, D.C. 20546  
NASA HEADQUARTERS





**NASA  
Reference  
Publication  
1120**

**1984**

# Solar Terrestrial Physics: Present and Future

*A Report Based on the Solar  
Terrestrial Physics Workshop,  
December 1982 to November 1983*

*Edited by*  
D. M. Butler and  
K. Papadopoulos

*University of Maryland  
College Park, Maryland*

Sponsored by the Office of Space Science  
and Applications, National Aeronautics  
and Space Administration and the  
National Science Foundation

**NASA**  
National Aeronautics  
and Space Administration  
**Scientific and Technical  
Information Branch**



## EDITOR'S PREFACE

The first workshop on Solar Terrestrial Physics began in December 1982 with the formation of twelve working groups covering critical subjects in solar terrestrial physics, and concluded with a meeting at the Coolfont Recreation Center, Berkeley Springs, West Virginia on June 6-10, 1983. The workshop sponsored by the Office of Space Science and Applications, National Aeronautics and Space Administration and the National Science Foundation assembled for the first time a community of scientists from various areas of solar terrestrial physics aiming at a unified view of the solar terrestrial environment as an interactive system. A steering committee composed of J. Boris, Naval Research Laboratory; D. Butler, NASA Headquarters; M. Kundu, University of Maryland; D. Papadopoulos, University of Maryland (Chair); E. Szuszcwicz, Naval Research Laboratory; and D. Williams, John Hopkins Applied Physics Laboratory appointed subcommittees on the areas of Ionospheric Physics, Magnetospheric Physics, Interplanetary Physics, and Solar Physics. With the advice of the subcommittees, the steering committee proceeded to select the subjects for the twelve working groups and invited the participants. The subcommittee and working group membership is listed in Appendix A.

The goals of the workshop were threefold. First, the diverse segments of the solar terrestrial community needed to be brought together. Represented were solar, solar wind, magnetospheric, ionospheric, and laboratory plasma physicists. This composition not only allowed for an overall view of the solar terrestrial system as an interactive system, but also improved the communication of ideas and the cross fertilization on subjects which share a common physical basis. Joint sessions among the various working groups strongly contributed towards this goal. Two working group subjects (reconnection and acceleration) are an integral part of a variety of solar terrestrial physics areas and naturally contributed to an integrated prospective.

Second, we desired a critical assessment of the current understanding of the subject by comparing in detail theoretical models with experimental observations. This provided for an opportunity to focus on the important scientific issues and elucidate the type of research work and methodology required for their resolution.

Third, a document assessing the state of understanding and underlining the important future issues in solar terrestrial physics was needed. The document was aimed at the level of the first or second year graduate student entering the field in need of an overview and an exposition of the critical issues of the subject.

The workshop was organized so as to facilitate these goals. The working groups developed and circulated draft papers on their subjects prior to the Coolfont meeting. These papers formed the basis of the discussions of the plenary sessions as well as the individual and joint working group sessions during the meeting. A draft report was completed for each working group shortly after the end of the meeting. The present manuscript "Solar Terrestrial Physics - Present and Future" is composed of a slightly revised version of the Coolfont report which took into account comments from a more general community during the period between July 1 and November 30, 1983.

We believe that the workshop accomplished its goal. Certainly, a major portion of the solar terrestrial physics community participated and was able to meet, discuss, criticize, expose and be exposed to the various topics. The informality of the workshop and the rather unstructured agenda allowed for many impromptu sessions and discussions lasting till the early morning hours and hopefully provided a major stimulus for new research inroads.

The present document was written as a beginning rather than an end. Indeed, if it serves its purpose well, it will be used a great deal in the current decade and not much thereafter. Solar terrestrial physics is a rapidly developing field, and it is expected that the upcoming experimental programs coupled with major advances in theory and simulations will explosively move the frontiers of our research. We hope that the utility of the document will affect areas much beyond solar terrestrial physics, such as laboratory plasma physics, planetary and cometary physics, and astrophysics. All of the above areas share many commonalities in the physical processes that control and manifest their behavior. A fundamental factor in the progress of understanding remotely probed plasma objects, such as astrophysical plasmas, is due to understanding the analogous processes operating in the solar terrestrial environment which has been diagnosed in situ and modelled by using sophisticated plasma physics theory and computations. As will become obvious to the reader, whether one is examining global problems such as the magnetospheric substorms or the coupling of the solar wind to the magnetosphere and of the magnetosphere to the ionosphere, or more local problems such as coronal transients, collisionless shocks or ionospheric structure at high latitudes, similar physics issues recur. These are magnetic field reconnection, particle acceleration, particle energy, and momentum confinement and transport across separatrixes, and plasma heating and radioemission due to the collective collisionless or collisional interactions. All of these issues are dealt with in the report either as individual topics stressing the basic physics and its applicability to various regions of the solar terrestrial environment or as components of the global subjects.

The number of people that contributed to the success of the workshop is so large as to preclude individual acknowledgements. We would like, however, to express our appreciation to the Office of Space Science and Applications, NASA, and the Atmospheric Division of the National Science Foundation for providing financial support as well as moral encouragement during the workshop. The staff of the Astronomy Program, the University of Maryland, and the Systematics General Corporation performed a commendable job in providing logistics and administrative support. The outstanding and dedicated work by the members of the steering committee and subject subcommittees is greatly appreciated. Collectively, the greatest thanks must go to the working group participants with special mention for the working group chairmen for accomplishing their enormous task with maximum efficiency. It has been a pleasure to associate with them.

D. M. Butler  
K. Papadopoulos

# TABLE OF CONTENTS

<u>SECTION</u>	<u>PAGE</u>
PREFACE .....	iii
EXECUTIVE SUMMARY .....	1
CHAPTER 1	
RECONNECTION OF MAGNETIC FIELDS	
I. INTRODUCTION .....	1-3
II. RECONNECTION AT THE MAGNETOPAUSE .....	1-7
QUASI-STEADY MAGNETOPAUSE RECONNECTION .....	1-8
FLUX TRANSFER EVENTS .....	1-10
III. RECONNECTION IN THE GEOMAGNETIC TAIL .....	1-11
IV. PARTICLE ACCELERATION IN CURRENT SHEETS .....	1-17
V. THE DIFFUSION REGION .....	1-20
DIRECTIONS FOR FUTURE WORK .....	1-25
VI. LABORATORY SIMULATION OF MAGNETOSPHERIC RECONNECTION .....	1-26
EXPERIMENTAL RESULTS .....	1-26
DISCUSSION .....	1-30
VII. RECONNECTION AND THE EVOLUTION OF SOLAR FIELDS .....	1-31
BIPOLAR MAGNETIC REGIONS .....	1-32
CORONAL HOLE BOUNDARIES .....	1-33
MAGNETIC FLUX DISAPPEARANCE .....	1-34
DISCUSSION .....	1-35
VIII. RECONNECTION IN SOLAR FLARES .....	1-35
SUMMARY .....	1-39
IX. ACCELERATION OF SOLAR-FLARE ELECTRONS BY RECONNECTION .....	1-39

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
X. MHD SIMULATION AND SMALL SCALE STRUCTURE .....	1-40
TURBULENCE IN RECONNECTION STUDIES .....	1-40
RECONNECTION SIMULATIONS USING THE SPECTRAL METHOD .....	1-42
CONCLUSION .....	1-43
XI. LABORATORY OBSERVATIONS OF NEUTRAL SHEETS .....	1-45
CURRENT DISRUPTION .....	1-46
MICROINSTABILITIES .....	1-47
ELECTRON DISTRIBUTION FUNCTIONS .....	1-47
XII. RECONNECTION IN TOKAMAKS .....	1-52
TWO-DIMENSIONAL RECONNECTION RATES .....	1-52
THREE-DIMENSIONAL SIMULATION OF TOKAMAK DISCHARGES .....	1-57
THREE-DIMENSIONAL RECONNECTION RATES .....	1-59
SUMMARY .....	1-60
XIII. CONCLUSION .....	1-60
XIV. REFERENCES .....	1-62

CHAPTER 2  
PARTICLE ACCELERATION

I. INTRODUCTION .....	2-3
II. PHENOMENOLOGY .....	2-3
SOLAR FLARES .....	2-5
ACCELERATION PROCESSES IN THE INTERPLANETARY MEDIUM .....	2-8
Interplanetary Shock Acceleration .....	2-10
Corotating Ion Streams .....	2-11
The Anomalous Cosmic Ray Component .....	2-12
Ions Upstream of Planetary Bowshocks .....	2-12
MAGNETOSPHERIC AND IONOSPHERIC OBSERVATIONS .....	2-13
Auroral Particle Acceleration .....	2-13
Magnetotail .....	2-16
Bowshock .....	2-17
PARTICLE ACCELERATION OUTSIDE THE SOLAR SYSTEM .....	2-19

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
III. THEORETICAL PARTICLE ACCELERATION MECHANISMS .....	2-20
ADIABATIC COMPRESSION, MAGNETIC PUMPING, AND DIFFUSION .....	2-20
ACCELERATION IN DIRECT ELECTRIC FIELDS .....	2-22
Current and Neutral Sheet Particle Acceleration .....	2-22
Particle Acceleration by Parallel Electric Fields .....	2-26
STOCHASTIC ACCELERATION .....	2-29
SHOCK PARTICLE ACCELERATION .....	2-32
Shock Drift Mechanism .....	2-32
Compression Mechanism .....	2-33
Lower Hybrid Wave Acceleration at Shocks .....	2-34
COHERENT WAVE ACCELERATION .....	2-35
INJECTION .....	2-36
IV. THE REMAINING QUESTIONS .....	2-37
SOLAR PARTICLE ACCELERATION .....	2-37
INTERPLANETARY PARTICLE ACCELERATION .....	2-37
MAGNETOSPHERIC PARTICLE ACCELERATION .....	2-38
Auroral Particle Acceleration .....	2-38
Magnetotail Particle Acceleration .....	2-38
Bowshock Particle Acceleration .....	2-38
EXTRASOLAR PARTICLE ACCELERATION .....	2-39
THEORETICAL MODELS FOR PARTICLE ACCELERATION .....	2-39
Adiabatic Compression and Diffusion .....	2-39
Direct Electric Field Acceleration .....	2-39
Shock Acceleration .....	2-39
V. REFERENCES .....	2-40

CHAPTER 3  
EVOLUTION OF SOLAR MAGNETIC FLUX

I. INTRODUCTION .....	3-3
II. APPEARANCE OF MAGNETIC FLUX .....	3-5
MODELS FOR FLUX EMERGENCE .....	3-5
UNEXPLAINED OBSERVATIONS .....	3-10
Small-Scale Flux Emergence .....	3-10
Anticorrelation of Small- and Large-Scale Emergence .....	3-11
Organization in Large-Scale Emergence .....	3-11

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
III. DYNAMICS OF SURFACE MAGNETIC FLUX .....	3-13
MAGNETIC FLUX TRANSPORT .....	3-13
Physical Processes Which Transport Flux .....	3-13
A Model for Flux Transport .....	3-15
MAGNETIC FLUX STRUCTURE .....	3-16
Flux Concentration .....	3-16
Weak Background Fields .....	3-17
IV. DISAPPEARANCE OF MAGNETIC FLUX .....	3-18
THEORETICAL CONSIDERATIONS .....	3-18
OBSERVATIONS OF FLUX DISAPPEARANCE .....	3-21
Small-Scale Disappearance of Bipolar Loops .....	3-21
Sunspot Erosion .....	3-21
V. SUMMARY .....	3-23
VI. REFERENCES .....	3-29

CHAPTER 4  
MHD WAVES AND TURBULENCE IN THE SUN  
AND INTERPLANETARY MEDIUM

I. INTRODUCTION .....	4-3
II. GLOBAL OSCILLATIONS OF THE SUN .....	4-4
III. OBSERVATIONS RELATED TO WAVES OR TURBULENCE IN THE SOLAR ATMOSPHERE .....	4-5
CHROMOSPHERE, TRANSITION REGION AND INNER CORONA .....	4-5
SOLAR WIND ACCELERATION REGION .....	4-7
IV. LOCAL WAVES IN THE SOLAR ATMOSPHERE: THEORETICAL CONSIDERATIONS .....	4-9
V. INTERPLANETARY HYDROMAGNETIC FLUCTUATIONS .....	4-12
VI. RECENT STUDIES OF THE INTERPLANETARY PLASMA BASED ON TURBULENCE THEORY .....	4-13
VII. EFFECTS OF WAVES AND TURBULENCE ON THE SOLAR WIND .....	4-17
VIII. REFERENCES .....	4-18



## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
CHAPTER 5	
COUPLING OF THE SOLAR WIND TO THE MAGNETOSPHERE	
I. INTRODUCTION .....	5-3
MAGNETOSPHERIC PLASMA REGIMES AND BOUNDARIES .....	5-3
MAJOR QUESTIONS CONCERNING SOLAR WIND-MAGNETOSPHERE COUPLING .....	5-6
II. MAJOR ELEMENTS OF SOLAR WIND-MAGNETOSPHERE COUPLING .....	5-6
THE ENTRY EFFICIENCY OF SOLAR WIND PLASMA .....	5-7
RECONNECTION OF GEOMAGNETIC AND INTERPLANETARY MAGNETIC FIELDS .....	5-7
CONTROL OF DAYSIDE CONVECTION PATTERNS	
BY THE EAST-WEST IMF COMPONENT .....	5-8
MAGNETIC TOPOLOGY OF THE LOW-ALTITUDE BOUNDARY LAYER .....	5-8
CORRELATION OF GEOMAGNETIC ACTIVITY WITH SOUTHWARD IMF COMPONENTS ....	5-9
SUNWARD CONVECTION IN THE POLAR CAPS .....	5-10
FLUX TRANSFER EVENTS AT THE DAYSIDE MAGNETOPAUSE .....	5-10
III. SOLAR WIND COUPLING MECHANISMS .....	5-11
PROPOSED COUPLING PROCESSES .....	5-11
MOMENTUM AND ENERGY TRANSFER .....	5-13
BOUNDARY CONDITIONS .....	5-16
IV. GLOBAL MODELING .....	5-18
V. SUMMARY AND OUTSTANDING PROBLEMS .....	5-20
VI. REFERENCES .....	5-22
CHAPTER 6	
CORONAL TRANSIENTS AND THEIR INTERPLANETARY EFFECTS	
I. INTRODUCTION .....	6-3
II. BACKGROUND MATERIAL .....	6-4
ANCIENT HISTORY--SOLAR FLARES AND GEOMAGNETIC STORMS .....	6-4
MODERN HISTORY--INTERPLANETARY SHOCK WAVES .....	6-5
CORONAL TRANSIENTS OR MASS EJECTIONS .....	6-8
III. THE PRESENT .....	6-10
THEORETICAL MODELS .....	6-10
NEW OBSERVATIONS OF CORONAL MASS EJECTIONS .....	6-13

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
Comparison of Mass Ejections seen by Skylab, SMM, and SOLWIND Coronagraphs .....	6-14
Mass Ejections in the Low Corona .....	6-15
Mass Ejections in the Outer Corona .....	6-15
Associations Between Coronal Mass Ejections and Interplanetary Phenomena .....	6-17
Associations Between Coronal Mass Ejections and the Phenomena Revealed by Radio Observations .....	6-19
IV. THE FUTURE .....	6-20
SOLAR ORIGINS .....	6-21
INTERPLANETARY EFFECTS .....	6-23
V. REFERENCES .....	6-24

CHAPTER 7  
CONNECTION BETWEEN THE MAGNETOSPHERE AND IONOSPHERE

I. INTRODUCTION .....	7-3
II. GLOBAL SCALE COUPLING .....	7-4
MAGNETOSPHERIC CONFIGURATION EFFECTS .....	7-4
ELECTRICAL COUPLING .....	7-7
PLASMA TRANSPORT .....	7-10
Solar Wind and Plasma Sheet .....	7-10
Ionospheric Sources .....	7-15
III. INTERMEDIATE SCALE COUPLING .....	7-18
EVENING POTENTIAL STRUCTURE .....	7-18
ELECTRIC FIELD SIGNATURES .....	7-22
MORNINGSIDE STRUCTURES .....	7-26
WESTWARD TRAVELING SURGE .....	7-28
IV. MICROSCOPIC SCALE EFFECTS .....	7-29
AURORAL KILOMETRIC RADIATION .....	7-29
ION CONICS AND BEAMS .....	7-32
V. SUMMARY/CONCLUSIONS .....	7-34
VI. REFERENCES .....	7-36

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
CHAPTER 8 SUBSTORMS IN THE MAGNETOSPHERE	
I. INTRODUCTION .....	8-3
II. SOLAR WIND - MAGNETOSPHERE COUPLING .....	8-7
INTERPLANETARY PARAMETER CHANGES .....	8-7
"BEST" SUBSTORM PREDICTORS IN THE SOLAR WIND .....	8-7
TIME DELAYS: SOLAR WIND - MAGNETOSPHERE COUPLING RESPONSE TIME .....	8-10
III. DRIVEN PROCESSES IN SUBSTORMS .....	8-11
DISSIPATION .....	8-11
Joule Heating .....	8-14
Particle Precipitation .....	8-15
Ring Current Injection .....	8-16
STORAGE .....	8-17
Tail Field Enhancement .....	8-17
Plasma Sheet Changes .....	8-17
IV. THE UNLOADING PROCESS IN SUBSTORMS .....	8-18
ONSET .....	8-18
Internal Trigger .....	8-18
External Trigger .....	8-20
Feedback .....	8-21
Location and Localization .....	8-21
RELEASE .....	8-22
A Model for Energy Release .....	8-22
Plasma Sheet Thinning and Recovery .....	8-26
$B_z$ Polarity .....	8-26
Plasma Flow .....	8-26
Open Field Line Identification .....	8-27
Some Statistical Results .....	8-27
Impulsive Plasma Injections .....	8-29
Energetic Particle Acceleration .....	8-34
The Substorm Current Wedge .....	8-37

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
RECOVERY .....	8-41
Possible Causes .....	8-41
Quenching and Multiple Onsets .....	8-41
V. SUMMARY AND FUTURE DIRECTIONS .....	8-42
VI. REFERENCES .....	8-45

CHAPTER 9  
IMPACT OF FLARES ON THE SOLAR TERRESTRIAL ENVIRONMENT

<u>SECTION</u>	<u>PAGE</u>
I. INTRODUCTION .....	9-3
II. THE FLARE PROCESS .....	9-3
THE IMPULSIVE PHASE .....	9-4
SUBSEQUENT ACCELERATION .....	9-8
OUTSTANDING QUESTIONS .....	9-10
III. IMPACTS ON THE HELIOSPHERE .....	9-11
SOLAR WIND HEATING .....	9-12
SHOCK ACCELERATION .....	9-12
FORBUSH DECREASES .....	9-14
GEOMAGNETIC ACTIVITY .....	9-14
OUTSTANDING QUESTIONS .....	9-17
IV. IMPACTS ON THE MAGNETOSPHERE .....	9-17
ENERGETIC PARTICLES .....	9-18
SOLAR WIND EFFECTS .....	9-18
OUTSTANDING QUESTIONS .....	9-19
V. IMPACTS ON THE IONOSPHERE AND ATMOSPHERE .....	9-20
X-RAY AND EUV EFFECTS .....	9-20
SOLAR PARTICLE EFFECTS .....	9-21
MAGNETOSPHERIC STORM EFFECTS .....	9-23
EFFECTS ON RADIO WAVE PROPAGATION .....	9-23
NEUTRAL ATMOSPHERE EFFECTS .....	9-24
ATMOSPHERIC ELECTRICITY EFFECTS .....	9-25
OUTSTANDING QUESTIONS .....	9-26

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
VI. IMPACTS ON TECHNOLOGY .....	9-27
RADIO COMMUNICATIONS .....	9-27
SCINTILLATIONS OF SATELLITE SIGNALS .....	9-27
CABLE COMMUNICATIONS .....	9-28
GROUND-BASED POWER SYSTEMS .....	9-29
PIPELINE .....	9-29
THE SPACE RADIATION ENVIRONMENT AND SPACE SYSTEMS .....	9-29
PEOPLE IN SPACE .....	9-31
SPACECRAFT GLOW .....	9-31
OUTSTANDING QUESTIONS .....	9-32
VII. SUMMARY .....	9-32
VIII. REFERENCES .....	9-33

## CHAPTER 10

## COLLISIONLESS SHOCK WAVES IN THE SOLAR TERRESTRIAL ENVIRONMENT

I. INTRODUCTION .....	10-3
II. NATURAL SHOCKS .....	10-4
OCCURRENCE .....	10-4
GLOBAL FEATURES .....	10-5
MHD Configurations .....	10-5
Shock Extensions and Global Asymmetries .....	10-9
Local Macroscopic Shock Profiles .....	10-10
KEY QUESTIONS .....	10-15
III. QUASI-PERPENDICULAR, SUPERCRITICAL PROCESSES .....	10-16
ION DYNAMICS AND FIELD STRUCTURE .....	10-17
Laboratory-Shock Structure .....	10-17
Natural Bowshock of the Earth .....	10-18
Numerical Simulations .....	10-19
Comparison of Simulation with the Observations of 7 November 77 .....	10-21
Ion Reflection Machinery .....	10-23
ELECTRONS .....	10-25
INSTABILITIES AND WAVES .....	10-26
IV. QUASI-PERPENDICULAR SUBCRITICAL PROCESSES .....	10-30

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
V. QUASI-PARALLEL PROCESSES .....	10-31
TERMINOLOGY .....	10-32
ION PARTICIPATION .....	10-32
MINGLED PROCESSES .....	10-33
PHENOMENOLOGY .....	10-33
VI. ION ACCELERATION .....	10-35
THE FORESHOCK .....	10-35
PROTON INJECTION INTO THE UPSTREAM FLOW .....	10-37
ION ACCELERATION AND HYDROMAGNETIC WAVE EXCITATION .....	10-40
VII. OUTSTANDING PROBLEMS .....	10-43
VIII. INVESTIGATE AVENUES .....	10-44
LABORATORY EXPERIMENTS .....	10-44
THEORY AND SIMULATION .....	10-45
SPACECRAFT OBSERVATIONS .....	10-46
IX. REFERENCES .....	10-50
CHAPTER 11	
ASSESSMENT OF PLASMA TRANSPORT AND CONVECTION AT HIGH LATITUDES	
I. INTRODUCTION .....	11-3
II. CURRENTS AND CONDUCTIVITIES .....	11-4
MEASUREMENT TECHNIQUES .....	11-6
THEORETICAL APPROACHES .....	11-6
CONDUCTIVITIES .....	11-7
CURRENTS .....	11-7
MAJOR UNRESOLVED PROBLEMS .....	11-8
DIRECTION OF FUTURE RESEARCH .....	11-8
III. THERMOSPHERIC WINDS AND PLASMA TRANSPORT .....	11-8
IONOSPHERE - THERMOSPHERE COUPLING .....	11-8
COMPARISON OF THEORY AND OBSERVATIONS .....	11-11
MAJOR UNRESOLVED PROBLEMS .....	11-11
DIRECTION OF FUTURE RESEARCH .....	11-13

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
IV. PLASMA CONVECTION .....	11-13
AVERAGE CONVECTION MODELS .....	11-14
TWO-DIMENSIONAL STRUCTURE OF CONVECTION .....	11-16
DYNAMIC VARIATIONS OF CONVECTION .....	11-17
IONOSPHERIC CONVECTION TIMES .....	11-17
IONOSPHERIC RESPONSE TO CONVECTION .....	11-20
DIRECTION OF FUTURE RESEARCH .....	11-27
V. POLAR WIND .....	11-27
THEORETICAL MODELS .....	11-28
OBSERVATIONS .....	11-30
DIRECTION OF FUTURE RESEARCH .....	11-31
VI. CONCLUSIONS .....	11-31
VII. REFERENCES .....	11-32

CHAPTER 12  
HIGH LATITUDE IONOSPHERIC STRUCTURE

I. INTRODUCTION .....	12-3
IONOSPHERIC STRUCTURE IN GENERAL .....	12-3
EQUATORIAL SPREAD-F IRREGULARITIES - A SUCCESS STORY AND A GUIDE FOR HIGH LATITUDES .....	12-4
FOCUS ON HIGH LATITUDE STRUCTURE .....	12-5
II. SOURCES AND OBSERVATIONS OF HIGH LATITUDE STRUCTURE .....	12-7
ELECTRON PRECIPITATION STRUCTURES .....	12-7
ELECTRIC FIELDS .....	12-11
FIELD-ALIGNED CURRENTS .....	12-13
PLASMA DENSITY STRUCTURE .....	12-13
III. PLASMA INSTABILITY THEORY .....	12-18
MACROINSTABILITIES AND HIGH LATITUDE STRUCTURE .....	12-18
F-Region Irregularities .....	12-18
E-Region Irregularities .....	12-25
MICROINSTABILITIES AND HIGH LATITUDE STRUCTURE .....	12-26
General Comments .....	12-26
Linear Theories .....	12-26
Nonlinear Effects .....	12-29

## TABLE OF CONTENTS (Continued)

<u>SECTION</u>	<u>PAGE</u>
IV. AN EMERGING PICTURE .....	12-29
V. FUTURE STUDIES .....	12-31
THEORETICAL THRUSTS .....	12-31
EXPERIMENTAL EMPHASIS .....	12-32
VI. REFERENCES .....	12-33
APPENDIX A - STEERING COMMITTEE AND SUBCOMMITTEE MEMBERS .....	A-1
APPENDIX B - WORKING GROUP PARTICIPANTS AND ATTENDEES .....	B-1
APPENDIX C - SUMMARY OF IONOSPHERIC PHYSICS .....	C-1



# LIST OF ILLUSTRATIONS

## CHAPTER 1 RECONNECTION OF MAGNETIC FIELDS

<u>FIGURE</u>		<u>PAGE</u>
1-1	Evolution of Magnetic Field Configuration in a Laboratory Reconnection Experiment .....	1-4
1-2	Two-Dimensional Steady MHD Reconnection Models .....	1-6
1-3	Tearing Mode Magnetic Field Configuration .....	1-7
1-4	Meridional View of the Magnetosphere Undergoing Steady Reconnection .....	1-8
1-5	Northern Half of a Flux Transfer .....	1-11
1-6	Computer-Simulated Evolution of Magnetic Fieldlines .....	1-14
1-7	Computer-Simulated Evolution of the Velocity Field and the Magnetic Neutral Line .....	1-15
1-8	Computer-Simulated Electric Field Vectors and Magnetic Neutral Line .....	1-16
1-9	Particle Motion and Non-Adiabatic "Speiser" Ion Motion .....	1-18
1-10	Field-Aligned Accelerated Ion Beams on the Outer Boundary of the Plasma Sheet During Substorm Recovery .....	1-21
1-11	Lower-Hybrid-Drift Instability Mode Structure .....	1-23
1-12	Time Evolution of the Current Density and the Magnetic Field in the Presence of a Model Anomalous Resistivity .....	1-24
1-13	UCR T-1 Eleven Meter Terrella Device Schematic Drawing .....	1-27
1-14	T-1 Magnetic Map at Alfvén Mach Number 2.5 .....	1-28
1-15	Magnetic Field Distribution on the Dayside of the Magnetosphere with a Southward DMF for Alfvén Mach Number 10 .....	1-29
1-16	Possible Birth by Magnetic Reconnection of a Transequatorial Loop .....	1-32
1-17	Coronal Holes Observed by Skylab .....	1-34
1-18	Magnetic Field Lines and Flow Vectors During a Line-Tied Reconnection Experiment .....	1-36
1-19	Height and Velocity of Loops During a Major Flare .....	1-38
1-20	Magnetic Field Lines, Fluid Streamlines, Constant Electric Current Density Contours, and Constant Vorticity Contours after 21.4 Alfvén Transit Times of Evolution .....	1-44
1-21	UCLA Reconnection Experiment Schematic .....	1-45
1-22	Frequency and Wavenumber Spectra of Magnetic Fluctuations in the Neutral Sheet .....	1-48
1-23	Measured Distribution Function Displayed as Surfaces of Constant F .....	1-49
1-24	Spatial Variation of the Distribution Function in the Neutral Sheet .....	1-50
1-25	Two-Probe Cross-Correlation Measurements of Temperature Fluctuations .....	1-51
1-26	Time History of Magnetic Surfaces in an "m=1 flip" Process .....	1-53
1-27	Profiles of Plasma State during the Reconnection Phase .....	1-54
1-28	Numerical Connection Rate Compared to Modified Sweet-Parker Scaling .....	1-55
1-29	Profiles of a Plasma State During a Nonlinear Evolution of the m=2 "Purely Resistive" Instability .....	1-56
1-30	Experimental Soft X-Ray Signals (BO Channel) from a PDX Tokamak Discharge .....	1-57
1-31	Details of Experimental Soft X-Ray Signals (BO Channel) .....	1-58
1-32	Poincare Map of Magnetic Field Lines .....	1-59

## LIST OF ILLUSTRATIONS (Continued)

CHAPTER 2  
PARTICLE ACCELERATION

<u>FIGURE</u>		<u>PAGE</u>
2-1	Schematic Picture of Interaction Between Two Solar Coronal Magnetic "Loops" .....	2-6
2-2	Schematic Representation of Acceleration Processes Observed in the Interplanetary Medium .....	2-9
2-3	Schematic Summary of Important Plasma Populations Encountered by a Low Altitude Satellite in the Auroral Region .....	2-14
2-4	Particle Acceleration Geometry in the Terrestrial Magnetotail .....	2-16
2-5	Schematic Illustration of Non-Uniform Bowshock Structure in Average Solar Wind with Magnetic Field at Average Stream Angle .....	2-18
2-6	Geometry Appropriate for Ion Acceleration by Electric Fields Perpendicular to the Plane Containing Magnetic Field Lines .....	2-25
2-7	Magnetic Field and Flow Configuration in a Shock: Three Views .....	2-33

CHAPTER 3  
EVOLUTION OF SOLAR MAGNETIC FLUX

3-1	Four X-Ray Photographs and a Magnetogram of a Pair of Active Regions .....	3-7
3-2	Three Rotations of a Long-Lived Active Region .....	3-8
3-3	Four Composite Images of an Emerging Flux Region in the Photosphere .....	3-12
3-4	Mechanisms of Flux Disappearance from the Photosphere .....	3-20
3-5	Time-Lapse Sequence of Fe XV Coronal Spectroheliograms and Photosphere Magnetograms .....	3-22
3-6	A Schematic Flux Tube Life Cycle: The Solar Clock .....	3-27
3-7	Geometry of Field Lines Beneath the Solar Surface .....	3-28

CHAPTER 4  
MHD WAVES AND TURBULENCE IN THE SUN AND INTERPLANETARY MEDIUM

4-1	Microturbulent Velocity as a Function of Temperature in the Solar Atmosphere .....	4-6
-----	--	-----

CHAPTER 5  
COUPLING OF THE SOLAR WIND TO THE MAGNETOSPHERE

5-1	Schematic View of Plasma Transport in Magnetospheric Boundary Layers .....	5-4
5-2	Magnetotail Cross-Section Sketch of the Major Magnetospheric Plasma Regimes .....	5-5
5-3	Schematic Description of the Origin, Transport, Acceleration, and Storage of Magnetospheric Plasma .....	5-5
5-4	Simple View of an Open Magnetosphere .....	5-12
5-5	Schematic View of Flux Transfer Events Compared to Solar Wind Magnetosphere Coupling .....	5-14
5-6	Field-Aligned Currents Generated in the Boundary Regions .....	5-16

## LIST OF ILLUSTRATIONS (Continued)

CHAPTER 6  
CORONAL TRANSIENTS AND THEIR INTERPLANETARY EFFECTS

<u>FIGURE</u>	<u>PAGE</u>
6-1	A Synthesis of Interplanetary Shock Observations and Their Interpretations .....6-6
6-2	Hypothetical "Cut-Off" of an Arcade of Distended Magnetic Loops .....6-7
6-3	A Model of Flare-Associated Shock Phenomena in the Solar Corona .....6-12
6-4	Radial Speeds of Coronal Mass Ejections as a Function of Height in the Low Corona .....6-16
6-5	Height of the White-Light Loop as a Function of Time .....6-16
6-6	SOLWIND Coronagraph Images Showing Coronal Mass Ejections .....6-18
6-7	Coronal Mass Ejections Moving Directly Toward Earth .....6-20

CHAPTER 7  
CONNECTION BETWEEN THE MAGNETOSPHERE AND IONOSPHERE

7-1	Nominal Features of Magnetospheric-Ionospheric Coupling .....7-4
7-2	Observational and Interpretation Features of Theta Auroras .....7-6
7-3	Four Magnetospheric/Ionospheric Current Patterns .....7-9
7-4	Birkeland-Current Pattern: Observed vs. Theoretical .....7-10
7-5	Plasma Sheet Boundary Energy Spectra .....7-11
7-6	Electron Distribution Signatures: Observed and Calculated .....7-13
7-7	Electron Cyclotron Harmonic Emissions .....7-14
7-8	Sounding-Rocket Observations of Electron Energy Spectra .....7-15
7-9	Typical Distribution Function Contours of Ion Beam, Ion Conic, and Electrons in an Inverted-V Structure .....7-17
7-10	Observational Features of Electric Field and Plasma Associated with a Discrete Arc .....7-19
7-11	Precipitated Electron Energy Fluxes: Measured and Calculated .....7-21
7-12	Electrostatic Shocks During a Typical S3-3 Satellite Passage from Plasmasphere to the Polar Cap .....7-24
7-13	Perpendicular and Parallel AC Electric Field Components .....7-25
7-14	Resistivity from Electron Interaction Momentum .....7-27
7-15	Energy-Time Spectra of a Dawn-Dusk Pass of the S3-3 Satellite .....7-28
7-16	Observed Auroral Kilometric Radiation (AKR) .....7-30
7-17	Typical AKR Wideband Spectrogram .....7-31

CHAPTER 8  
SUBSTORMS IN THE MAGNETOSPHERE

8-1	A Summary of the Substorm Sequence .....8-4
8-2	Schematic Cross-Sectional View of the Magnetosphere .....8-6
8-3	Four-Day Internal Measurements of the Solar Wind-Magnetosphere Dynamo Power and the Total Energy Dissipation Rate of the Magnetosphere .....8-12
8-4	Simple Representation of the Equivalent Circuit of a Directly Driven Magnetospheric Model .....8-13

## LIST OF ILLUSTRATIONS (Continued)

<u>FIGURE</u>		<u>PAGE</u>
8-5	Schematic Drawing of Representative Field Lines in the Dawn-Dusk Meridian Plane .....	8-14
8-6	Thermal Electron and Ion Motions Near a Quasi-Neutral Sheet .....	8-19
8-7	Three Photographs of Comet Morehouse .....	8-22
8-8	Plasma Sheet Configuration Changes During a Substorm According to the Neutral Line Model .....	8-24
8-9	Superposed Epoch Analysis of Five Plasma Measurements from 16 Events .....	8-28
8-10	Schematic Drawing of a Hypothetical Formation of an Injection Boundary Near Synchronous Orbit .....	8-33
8-11	Schematic Drawing of a Comprehensive Model for the Morphology of Energetic Ion Enhancements .....	8-36
8-12	Schematic Drawing of Current Flow on the Nightside of the Inner Magnetosphere During a Substorm .....	8-38
8-13	Generation of a Polarization Electric Field and an Incomplete Cowling Channel in a Region of Enhanced Ionization .....	8-39
8-14	Spatial Distribution of Hall Conductivity, Horizontal Electric Field Vectors, and Ionosphere and Field-Aligned Current in the Region of Active Break-up Aurora Around Local Magnetic Midnight .....	8-39
8-15	Spatial Distribution of Hall Conductivity, Horizontal Electric Field Vectors, and Ionosphere and Field-Aligned Current in the Region of Active Break-up Aurora for a Westward Travelling Surge in the Early Evening Sector .....	8-40

## CHAPTER 9

## IMPACT OF FLARES ON THE SOLAR TERRESTRIAL ENVIRONMENT

9-1	Occurrence Frequency of Solar Proton Events Near Earth During Solar Cycles 19 and 20 .....	9-4
9-2	Hard X-Ray Time Profiles as Observed by the HXRBS Instrument on SMM .....	9-5
9-3	Ten X-Ray Contour Plots from the HXIS Instrument on SMM .....	9-5
9-4	Flare Kernels of 1980 May 21 .....	9-7
9-5	Counting Rate Versus Time for Various Energy Channels of the SMM $\gamma$ -Ray Spectrometer During $\gamma$ -Ray Line Flare .....	9-9
9-6	Illustration of the Separation of Energetic Storm Particle Events with Quasi-Parallel Shocks and Particle "Spikes" with Quasi-Perpendicular Shocks .....	9-13
9-7	Four Forbush Decreases Associated with the Passage of Interplanetary Shocks ....	9-15
9-8	Schematic of a Kinky, Inclined Heliospheric Current Sheet and Neighboring Magnetic Fields .....	9-16
9-9	Comparison of Typical Electron Density Profiles for a Large Flare with Those Typical of the Quiet Daytime Ionosphere .....	9-21
9-10	Ion Production Rate in the Lower Ionosphere Produced by a Large SPE, 4 August 1972 .....	9-22
9-11	Effects of the 4 August 1972 SPE on the High Altitude Stratospheric Ozone Layer .....	9-22
9-12	Effects of the 4 August 1972 SPE on the Vertical Component of the Atmospheric Electric Field at the High Altitudes .....	9-25
9-13	Output Voltage of the Power-Feed Equipment at the European End of a Transatlantic Cable .....	9-28
9-14	Comparison of the Earth Currents Measured Near Fairbanks and the Induced Currents Measured in the Alaskan Pipeline at the Chena Test Site .....	9-30

## LIST OF ILLUSTRATIONS (Continued)

CHAPTER 10  
COLLISIONLESS SHOCK WAVES IN THE SOLAR TERRESTRIAL ENVIRONMENT

<u>FIGURE</u>		<u>PAGE</u>
10-1	Angular Locations on a Mean Shock Surface of Selected Shock Encounters .....	10-5
10-2	Major Regions of the Solar Wind's Interaction with the Earth's Magnetosphere ...	10-7
10-3	Average Bowshock Surfaces at Venus, Earth, Mars, Jupiter, and Saturn .....	10-8
10-4	Schematic Representation of Four Classes of Upstream Ion Populations .....	10-11
10-5	Two Extremes of Shock Appearance, with Ramp Length $L$ .....	10-11
10-6	Sketches of Three Representative Shock Profiles .....	10-14
10-7	Three-Dimensional Representation of Regions of $M-\theta B_n$ -Beta Space That Have Been Intensively Studied to Date .....	10-16
10-8	Schematic Diagram of a Perpendicular Shock Wave .....	10-20
10-9	Illustrations of Supercritical Quasi-Perpendicular Shock .....	10-22
10-10	Profiles of the Simulation Results for $O$ , $rm$ , and $ry$ .....	10-24
10-11	Sketches of Two Typical Examples of the Observed Upstream to Downstream Change in Electron Distribution .....	10-27
10-12	Time Sequence of Averaged Plasma Wave Spectra Through the Shock Crossing of 7 November 1977 .....	10-29
10-13	Trajectories and Velocity Space Diagram of Specularly Reflected Ions for Different Orientations of Interplanetary Field $B$ with Respect to the Shock .....	10-36
10-14	Trajectories and Velocity Space Diagrams of Ions of Various Pitch Angles Returned to Solar Wind from the Source Surface of Reflected Ions .....	10-39
10-15	Conceptual Diagram of Foreshock Ion Fermi-Acceleration in One Plane .....	10-41
10-16	Relative Scales of Important Plasma Quantities, Shock Dimensions, and Spacecraft Parameters .....	10-47
10-17	Relative Percentage of Occurrences of Subcategories of Quasi-Perpendicular Shock Conditions Dependent on Mach Number and Beta .....	10-49

CHAPTER 11  
ASSESSMENT OF PLASMA TRANSPORT AND CONVECTION AT HIGH LATITUDES

11-1	Block Diagram Showing the Coupling Between the Thermosphere, Ionosphere, and Magnetosphere .....	11-3
11-2	Cone-Shaped Regions Formed by Birkeland Currents Flowing into and away from the Auroral Region .....	11-4
11-3	Polar Plots of the Calculated Circulation over the Southern Hemisphere .....	11-10
11-4	Ion Heating Efficiency in Log Pressure Coordinates for Solar Maximum and Solar Minimum Condition .....	11-12
11-5	Generalized Log Pressure Profiles of the Neutral Gas Heating Efficiency for Auroral Electron Fluxes .....	11-12
11-6	Comparison of Heppner's Results and Observations to Those of Milestone Hill for a "Toward" IMF Configuration .....	11-14
11-7	Ion Drift Velocities Observed Along a Single Orbit Compared to the AE Model Results .....	11-15
11-8	Maps of Polarization Electric Field for a Representative Evening Sector .....	11-18
11-9	Electrodynamics of a Westward Traveling Surge .....	11-19
11-10	Contours of Electric Potential Derived from Radar Ion Drift Data .....	11-20

## LIST OF ILLUSTRATIONS (Continued)

<u>FIGURE</u>		<u>PAGE</u>
11-11	Plasma Convection Trajectories in a Magnetic Reference Frame .....	11-21
11-12	O <sup>+</sup> Density at 300 km for Weak Convection in Winter .....	11-23
11-13	O <sup>+</sup> Density Contours at 300 km for Strong Convection in Winter .....	11-24
11-14	O <sup>+</sup> Density Contours at 300 km for Strong Convection in Summer .....	11-25
11-15	Contours of the Ion Temperature in a Magnetic Quasi-Inertial Reference Frame ..	11-26
11-16	Schematic Diagram of H <sup>+</sup> Velocity Distribution Variation by Altitude for Collisionless Supersonic H <sup>+</sup> Outflow .....	11-29

CHAPTER 12  
HIGH LATITUDE IONOSPHERIC STRUCTURE

12-1	Basic Equatorial Spread F Geometry .....	12-4
12-2	Model of Plasma Enhancement Used in Numerical Simulations .....	12-6
12-3	Phenomenological Source Term Map of High Latitude Ionospheric Structure .....	12-8
12-4	Example of Polar Electron Precipitation Profile .....	12-10
12-5	Typical Signature of High Latitude Convection Velocity with Southward IMF .....	12-12
12-6	Electric Fields Observed from a Southern Hemisphere Polar Cusp Crossing of the DE-2 Satellite .....	12-14
12-7	<u>In Situ</u> Plasma Density Structure Across the Northern High Latitude Ionosphere as Measured on Three Polar Passes of the S3-4 Satellite .....	12-15
12-8	Altitude/Latitude Variation of Electron Density in the Midnight-Sector Auroral Zone .....	12-17
12-9	Physical Mechanism of the F Region $\underline{E} \times \underline{B}$ Gradient Drift Instability .....	12-20
12-10	Linear and Non-Linear Evolution of the $\underline{E} \times \underline{B}$ Instability in a Convecting Auroral F-Region Plasma Enhancement .....	12-21
12-11	East-West Structure/Latitude Variation of Electron Density at Different Altitudes .....	12-22
12-12	Physical Mechanism of the Current Convective Instability .....	12-24

# EXECUTIVE SUMMARY

## CHAPTER 1 - RECONNECTION OF MAGNETIC FIELDS

The study of magnetic field reconnection and its possible occurrence and importance in the magnetosphere and on the Sun has drawn, not only upon direct observations and measurements, but also upon available theory, upon computer and laboratory simulations, and upon examination of the process as it occurs in tokamaks. It has led to the following overall assessment.

The reconnection model has served, and continues to serve, as a successful organizing concept in magnetospheric as well as in solar physics. In the former area, recent spacecraft observations have helped move our capability to the point where direct quantitative comparisons between the predictions of simple reconnection theory and observations have become possible. These comparisons provided impressive support for reconnection both in the geomagnetic tail and at the magnetopause and demonstrated that the theory itself must be developed further if it is to serve as an effective guide in the future. Global as well as local computer models will undoubtedly be a principal means to achieve this goal, since it appears that global boundary conditions as well as the presence of local microscopic processes play an important role in the overall dynamics of magnetospheric reconnection. The success of computer simulations in helping to establish the occurrence and properties of reconnection in tokamaks serves as a useful illustration of this point.

The tokamak studies also show that reconnection can proceed even in circumstances where the magnetic field along the separator is the dominant field. Thus, any field-aligned current sheet or filament in the magnetosphere or elsewhere is a potential site for reconnection.

One of the most remarkable recent observational results is that magnetopause reconnection, which is thought of as a driven process, rarely occurs in a steady state. Rather it manifests itself in a patchy and highly time-dependent fashion. The theoretical aspects of these so-called flux transfer events are poorly understood and new theory as well as computer and laboratory simulation would be desirable. Flux transfer events promise to provide a major mechanism for the transfer of mass, momentum, and energy across the magnetopause leading, in an average sense, to a turbulent model of magnetopause reconnection. These observations also suggest that some kind of threshold, in addition to a southward magnetosheath field, may exist for the onset of magnetopause reconnection and that the magnetopause, either by accident or by design, operates near the threshold. This description agrees in a qualitative way with laboratory terrella experiments which indicate the stable occurrence of closed as well as open front-side magnetopauses. These facts indicate that we have an outstanding opportunity to identify from spacecraft observations what are the critical onset and switch-off conditions for magnetopause reconnection. A similar opportunity exists for impulsive reconnection in the tail.

Recent three-dimensional time-dependent numerical simulations of magnetotail reconnection have confirmed the basic features of classical reconnection models. They have also duplicated a number of additional observed substorm-related signatures and they contain many new features which may form a basis for further detailed comparisons with observations. In particular, field-aligned currents and complex neutral-sheet structure form integral parts of the models.

We take particular notice of the incomplete state of observational and theoretical knowledge concerning the plasma processes near the separator line, in the diffusion region. Microinstabilities may play an important role in this region but, if so, it has not been established which instabilities are most important. Information concerning these processes may perhaps be obtained from laboratory experiments but it is not clear that such results can be translated directly to the magnetosphere application. The possibility also exists that in the tail and at the magnetopause

resistively provided by particle inertial and gyroeffects dominates. Since the processes in the diffusion region may be instrumental in determining the onset and switch-off criteria for reconnection, the development of a better understanding of these processes is essential.

In light of the absence of in situ measurements, and given the complicated nature of the solar magnetic field, it is not surprising that our knowledge of solar magnetic field reconnection is less developed than that of magnetospheric reconnection. Nevertheless, recent spacecraft missions such as Skylab and Solar Maximum Mission (SMM) have provided a new, extremely valuable, data base for systematic studies of the reconnection process itself in its solar setting. Such studies should be vigorously pursued. On the theoretical side, computer model of flares, and models that draw upon our experience with reconnection in the magnetotail and in tokamaks offer promise. However, it must be remembered that the solar environment has many important features not present elsewhere. The possibility of coronal heating by small-scale turbulent reconnection is an outstanding example of a problem where recent numerical simulations may provide crucial insights.

## CHAPTER 2 - PARTICLE ACCELERATION

Particle acceleration is one of the more widespread processes in the solar terrestrial and astrophysical plasma systems. Its universality and the extensive observational base of its radiation signatures allows the community to draw not only upon in situ cause and effect measurements in the ionosphere, the magnetosphere and the interplanetary medium, but also upon high energy particle and radiation spectra from the Sun, upon plasma theory and computer simulations, and upon a wealth of informations and theoretical speculations from astrophysical systems. The subject has significantly progressed since its beginnings in cosmic ray theory research, both in the observational and theoretical areas; perhaps most gratifying is the agreement between theory and observations in at least one case, particle acceleration in the quasi-perpendicular bow shock, which is both well observed and diagnosed, and theoretically modeled. Many acceleration aspects, however, remain unclear. This is more so, since a description of acceleration and interpretation of the observations involves not only the acceleration as a physical process by itself, but its operation within the larger space system (e.g., magnetotail, shocks, auroral arcs etc.) and the physics of the conversion of particle energy to radiation for remotely diagnosed systems. The report covers acceleration in the Sun, the interplanetary medium, the auroral zones, the magnetotail, and the bow shock in a fashion that preserves the unity and commonality of the acceleration as a physical process.

In the absence of in situ measurements the acceleration properties of solar particles are inferred by remote sensing of the radiation signatures and by direct measurements in space. Major impetus in the research of flare acceleration was provided by the advent of the Solar Maximum Mission (SMM) and Hinotori satellites and the Very Large Array (VLA) telescope, which produced spatially resolved two dimensional images of the radiating sources in the radio band and at hard X-rays ( $< 30$  keV) with short time resolution. Observations of  $\gamma$ -rays and high energy neutrons combined with theoretical calculations allowed us to extend the knowledge of particle spectra in the energy range between 1 MeV - 1 GeV. Measurements of particle spectra in space further complemented the radiation diagnostics by providing (particle spectra and especially) a better understanding of the elemental, charge-state, and isotopic composition of energetic ions ( $< 1$  to  $> 20$  MeV) in solar energetic particle events (SEP).

The observations gave an important test ground of the theoretical models without, however, providing sufficient constraints for reliable testing. Presently the theoretical models are called upon to predict features accessible at present or in the near future, to measurements; missing elements include the identification of the acceleration process and its location(s), but also calculation of spectra and their dynamic properties, charge/mass ratio dependences and prediction of



observable detailed diagnostic features. On the observation side there is a need for both charge state and particle spectra for escaping populations for both "normal" and anomalous composition SEP events. High angular resolution x and Y-ray measurements or multiple spacecraft observations are needed for localizing the acceleration and interaction sites. An important question requiring resolution is the need for two or multi step acceleration processes.

In the interplanetary medium the acceleration sites seem to be associated with shocks; these shocks can be driven by solar flares, planetary magnetospheres (e.g. bow shocks), or are remote from the Earth (coronal shocks, shocks in corotating regions, the solar wind terminal shock). Most of the experimental evidence is from near Earth satellites. More recently measurements of intensity gradients with deep space probes allowed the identification and location of acceleration sites beyond the Earth. Some shock acceleration form has been conjectured as the mechanism responsible for the particle energization. The major needs presently are on the observational side. Events with associated very energetic (10 - 100 keV) electrons are relatively rare; hence, detailed IPM measurements of this component are few in number. Furthermore, there is a general need for measurements of 1 AU particle events and their identification with solar transients. What is the seed population of energetic ions (> 30 keV) associated with ESP events? Resolution requires a time-resolved observations of several distinct ion particle distributions (as will be done by the ISPM spacecraft) and their anisotropies. Measurements of low ionization charge states and ion composition, as well as high mass resolution observations, are also of great interest; such measurements are particularly relevant to understanding the process(es) leading to the anomalous component (in particular, of the anomalous species, such as oxygen) and acceleration in corotating streams, and to resolving details of the acceleration process in the vicinity of corotating streams (vz., the importance of curved geometry in these shocks).

The terrestrial magnetosphere has provided a rather comprehensive evidence of many processes leading to the production of fast particles. The auroral zones have been rather well diagnosed and shown the existence of a downward accelerated electron population, an electron population trapped between magnetic mirror points and an electrostatic barrier at high altitude, and an upward accelerated ion population of ionospheric origin. Theoretical modeling is evolving steadily and with a high degree of sophistication. The outstanding problem in auroral acceleration is the mechanism driving the electric potential across L shells; it seems that modeling of the process might require kinetic models. Other critical questions relate to the process and the extent to which the potential drop is localized, the mechanism that generates the ion conics above the auroral ionosphere and the nature of the relationship between the cross-tail convection potential, the upward flux of ions and the field-aligned potential drop over the polar cap.

Recent magnetotail observations are revealing a general pattern of the energetic particle population. The lobe has low number density and includes moderately energetic ions of magnetospheric origin. Both the densities and the ion energies increase towards the central plasma sheet with earthward streaming of tens of keV energy ions, followed by counterstreaming ions beams and ion temperatures in the 1-5 keV range. The central sheet also includes .1-1 keV electrons and substantial e-s wave activity in the lower hybrid branch. Finally very energetic electron (few hundreds of keV) and proton (in excess of 1 MeV) bursts are often observed in the tail. The physics of the ion thermalization while merging with the plasma sheet, of the electron heating in the sheet and the origin of the energetic particle bursts are the outstanding acceleration problems in the magnetotail.

The comprehensive ISEE -1,-2,-3 particle measurements of the terrestrial bow shock, complemented by similar observation at Jupiter and Saturn formed an outstanding data base for the study of shock acceleration. Very mature and sophisticated theories were driven by these data with substantial success, especially on the quasi-perpendicular shock region. The outstanding theoretical problem remains the quantitative understanding of the quasi-parallel shock acceleration process.

Theoretical modeling is the umbrella under which the generic features of the various acceleration processes can be explored in a manner emphasizing the physics and identifying the commonalities. Despite great progress, in understanding the fundamental physics of several acceleration mechanisms and in modeling their observable consequences in regions of the solar terrestrial environment, several important questions remain to be solved. In acceleration by adiabatic compression and diffusion, the relevant diffusion mechanism and the associated diffusion coefficient are quite uncertain. Designing observational tests that distinguish among diffusion processes is an important task. The role of injection conditions and of the magnetic geometry on the final energy is a key uncertainty in parallel electric field acceleration. The same holds true for the conditions under which microinstabilities give rise to anomalous resistivity caused by density fluctuations or micro-double layers and how it affects runaway acceleration. Energization during reconnection is usually thought as due to inductive electric fields. The possibility for acceleration caused by the shocks required in the Petschek and Sonnerup reconnection models has not been explored. Understanding of stochastic as well as coherent acceleration by electrostatic waves is far from resolution. On the widely used shock acceleration the basic problems revolve about the issues of (a) the relationship between the microstructure of shocks and the ambient plasma conditions (including, in particular, conditions upstream from the discontinuity), and (b) how the accelerated particle themselves feed back to affect the (micro-)structure of the accelerating shock. These problems are particularly acute for quasi-parallel shocks, which are not well-understood (viz., why is it, for example, that quasi-linear theory seems to 'work', even though  $\delta B/B \approx 0(1)$  ?), and for high Mach number (collisionless) shocks (including relativistic shocks, which are likely to be important to particle acceleration in the vicinity of pulsars and other collapsed objects). This area is -- just as the problem of the substorm, or reconnection, and the related particle acceleration -- a major research problem in its own right, and is the subject of a companion chapter in this document.

### CHAPTER 3 - EVOLUTION OF SOLAR MAGNETIC FLUX

Hale discovered the association of intense magnetic fields with sunspots 75 years ago. Observation of the sunspots for hundreds of years had recorded not only the fairly regular 11-year sunspot cycles but also irregular behavior, such as the Maunder Minimum, a period of about 70 years from 1645 to 1715 when essentially no sunspots were seen. Hale's discovery, however, was the first hint of complex magnetic field activity on the Sun. Today we know that sunspots and the associated bipolar magnetic regions (BMRs) appear on the Sun in large-scale patterns, they evolve, interact, migrate across the solar surface and then disappear following the sunspot cycle. Since the Sun's average dipole field reverses polarity every 11 years, a complete magnetic cycle lasts 22 years. Ground-based observations still provide most of our data base on the sunspots and magnetic activity.

In the last 25 years steadily improving observations, many of which are obtained above the atmosphere, have shown that the dynamic evolution of solar magnetic fields controls most components of solar variability affecting the geophysical environment such as disruptions of global communications. Observational and theoretical advances have given us a better view of the dynamic, sometimes violent, MHD and plasma phenomena which control sunspots and the solar transition region, heat the corona, drive the solar wind, and power flares, bright points, and spicules. Most of the phenomena investigated by the other eleven working groups depend on solar magnetic field variability to drive them or to trigger subsequent mechanisms which drive them.

There are other important reasons why a better understanding of the evolution of the solar magnetic flux is desired. Because the surface features can be observed in some detail, the Sun is an important source of information on plasma and MHD phenomena which can also occur in the laboratory, in our atmosphere, and in space. The Sun's large space scales and correspondingly

longer timescales make possible observations which cannot be obtained on the Earth. The Sun is also a star and therefore provides a wealth of information for extrapolation to and interpretation of other stars.

There are other important reasons why a better understanding of the evolution of the solar magnetic flux is desired. Because the surface features can be observed in some detail, the Sun is an important source of information on plasma and MHD phenomena which can also occur in the laboratory, in our atmosphere, and in space. The Sun's large space scales and correspondingly longer timescales make possible observations which cannot be obtained on the Earth. The Sun is also a star and therefore provides a wealth of information for extrapolation to and interpretation of other stars.

To understand the evolution of solar magnetic flux we must answer the fundamental question: how does the solar dynamo work? Recent observations certainly influence models of the dynamo, but we have not been able to deduce the subsurface flows which control the generation of magnetic flux and presumably the entire solar cycle. The universality and importance of the scientific issues have attracted some of our best scientific minds and kept them busy for the better part of a century.

The first modern qualitative model of the operation of the solar activity cycle was provided by Babcock (1961). He considered the evolution of an axisymmetric dipole field of concentrated flux ropes subject to differential rotation. His model was able to account for the reversal of sunspot polarities from cycle to cycle, the drift of sunspot activity from mid-latitudes toward the equator during the cycle, and the reversal of the global dipole field near each sunspot maximum. Leighton (1969) quantified some aspects of the model by obtaining numerical solutions of the linear dynamo equations. These solutions exhibited behavior akin to that of the large-scale magnetic fields on the Sun, including fluctuations in period and amplitude in response to fluctuations in the rate of emergence of new flux.

Perhaps the most striking recent finding about magnetic activity cycles is other evidence of their intermittency and aperiodicity. Observations of stellar spectra over extended periods show that stars also exhibit magnetic cycles, and that some, in particular fast rotators, show aperiodic magnetic activity. It appears that intermittency and aperiodicity may be natural consequences of nonlinear dynamos. Calculations for model dynamos with interactions between the magnetic field and differential rotation have shown a transition from periodic to quasi-periodic to chaotic behavior as parameters of the dynamo change. They have also produced a convective flow which shows both differential rotation and a magnetic cycle.

Lacking a clear view of subphotospheric phenomena, this chapter concentrates on what can be deduced from observations by addressing four issues:

- (1) How does flux emerge on the Sun?
- (2) How does magnetic flux disappear?
- (3) What is seen at the limits of resolution?
- (4) What large-scale patterns are detected?

The answer to the fundamental question, how does the solar dynamo work, can only evolve from interpreting clues uncovered in research aimed at these more "visible" questions. Numerous aspects of the first two questions arise repeatedly in theoretical attempts to describe physical processes which occur during the life cycle of magnetic flux on the Sun. The last two questions are observational and arise from needed advances in our ability to resolve and measure small time- and space-scale structures and in our ability to extract weak, large-scale patterns from the

noisy background over a long time period. This chapter considers the Appearance of Magnetic Flux (Section II); The Dynamics of Surface Magnetic Flux (Section III); and The Disappearance of Magnetic Flux from the surface of the Sun (Section IV). In each section the authors first discuss some of the standard models which represent generally accepted views, and then consider observations which seem to defy explanation within the context of these standard models.

Progress on the observational side has been rapid but we still do not have enough detailed, simultaneously obtained information about small-scale structures and dynamics to distinguish among many possible interactions and configurations. Specifically, it is crucial to see into the subarcsecond regime (50-500 km) and to identify and record phenomena in seconds rather than minutes to understand the dynamics of emergence and disappearance of flux. Small structures (~100 km) and correspondingly short timescales (~10 s) in the solar plasma temperature, density, velocity, and vector magnetic fields seem to control much of the dynamics that we do see.

Theory and numerical modeling are being used to bridge this gap, making increased use of supercomputers and simulation techniques. Macroscopic observables are estimated computationally assuming certain mechanisms and configurations which have been postulated but which have not yet been verified by direct observation. This approach will bear more fruit as corresponding observational improvements are made.

Clearly one direction of future research will be to obtain adequate resolution in time and space of the elemental flux tubes as they emerge. Rare pictures under conditions of particularly good viewing yield marginal resolution of some 100-200 km structures. To study emergence, submergence, and reconnection quantitatively will require not only better spatial resolution but also the capability to obtain a whole series of high resolution shots reliably. Space-based telescopes and instrumentation meet these criteria in a way that ground-based observations cannot match.

While theory can lead observation in some areas, it necessarily lags behind in others. Our theoretical understanding does not even encompass some large-scale solar phenomena which have become well established in the last two or three decades. Magnetic activity and related transient phenomena seem to present particularly difficult theoretical problems. Examples of observations waiting for a quantitative theory are coronal heating, solar prominences, coronal transients, the equatorward migration of the sunspot belts from the beginning to the end of each solar cycle, the acceleration of the solar wind, and the solar dynamo itself.

The problem of magnetic flux emergence is receiving more attention with the realization that most of the energy available in the magnetic flux beneath the surface already has been liberated in expansion by the time the persistent magnetic field configurations above the surface have formed. These dynamic flux-emergence problems are very difficult to solve and there has been little incentive to do so since magnetic transients in the low corona have rarely been seen. Though proposed, for example, as flare triggers, or as the output of reconnection, fine-scale magnetic transients have defied all attempts to measure them quantitatively because of the small length and time scales involved. And while there exist many examples of the emergence of a bipolar flux loop through the solar surface, observation of the submergence or disappearance of such loops is generally ambiguous. Only recently a few puzzling cases of "unipolar" flux emergence or decay have been seen in which the other polarity footpoint is missing or hidden.

The second observational issue, detecting weak, large-scale patterns, affects our understanding both of flux disappearance and of the magnetic dynamo-solar cycle. The relationship between the solar cycle and large-scale magnetic flux dispersal across the solar disk resulting from the combined action of the Sun's differential rotation and random walk diffusion has been recognized for two decades. More recently, poleward meridional flows have been invoked to help

explain this large-scale motion of flux and the timing of the polar field reversal. Measurements of meridional flow are difficult, however, and require extraction of  $\sim 10$  m/s flows over a large area where km/s transient flows are present almost everywhere. Correlating these flows with magnetic fields near the Sun's poles is further complicated by limb effects which degrade observations and by magnetic fields which lie almost perpendicular to the line of sight.

With roughly two complete solar cycles of regular full-disk magnetic data now available, solar physicists are beginning to understand the evolution of the large-scale, long-lived magnetic patterns. These features must now be related to the large convection cells beneath the solar surface, and hence to the solar dynamo, which may be driven in part by this "giant granulation." Again, space-based observation will be the key to achieving the continual coverage in both area and duration that is needed to average out short period oscillations and transients effectively.

#### CHAPTER 4 - MHD WAVES AND TURBULENCE IN THE SUN AND INTERPLANETARY MEDIUM

The deliberations of the working group centered on the current state of knowledge concerning the existence, nature and dynamics of MHD waves and turbulence in the solar atmosphere and interplanetary medium. We first considered remote-sensing observations of global oscillations of the Sun, of nonthermal motions in the chromosphere, transition region, and corona, and their possible interpretations in terms of waves or turbulence. We next considered the region of solar wind acceleration; here coronal imaging, resonance-line spectrometry and radio techniques provide the relevant observations. We then considered fluctuations and discontinuities in the interplanetary medium itself, in which the primary data come from direct sampling of the plasma by spacecraft-borne magnetometers and plasma analyzers. We do not address microscopic plasma instabilities or waves/turbulence and there is no attempt to provide a comprehensive review of any of these topics; rather the goal has been to assess how successful past attempts to close theory and observation have been, and to obtain a sense of where future advances may lie.

The energy source of the solar wind is still not known. It is generally agreed that heating or acceleration must occur over an extended region, probably governed by MHD waves/turbulence. Indeed, several studies of radio signals passing through the corona are consistent with existence of a field of large-amplitude fluctuations extended 10-20 solar radii out from the Sun. Such a wave field would exert an outward force on the solar wind plasma, thus accelerating the wind. Dissipation of the waves would heat the plasma (note that observed interplanetary proton temperatures are much hotter than they would be if their adiabatic cooling were modified only by thermal contact with the electrons).

A number of wave-driven models of the solar wind have been developed: models postulating a magnetoacoustic wave field invoke wave dissipation by Landau damping; models postulating an Alfvénic wave field invoke dissipation by nonlinear processes. Both kinds of models are readily tuned to give good agreement with interplanetary observations. Certainly both processes could be acting simultaneously. Observations of the flow and fluctuations within a few tens of solar radii will be necessary to determine which, if any, wave processes are operative. Optical and radio techniques (summarized earlier in this report) are valuable probes of this near-solar region that have only begun to be applied. However, these techniques involve integration over the line of sight, and the associated ambiguities of interpretation may not admit definitive evaluation of wave-driven models. Resolution of this question probably awaits in situ measurements from a space probe in the near solar region.

The wave-driven models also are incomplete in principle. The required amplitudes are large enough that nonlinear effects may be important (indeed the dissipation of Alfvén waves would have to be nonlinear). The physics of large-amplitude hydromagnetic waves in collisionless plasma is only partly understood, and this area is an important element of future progress. If the interplanetary fluctuations are in fact the remnant of fully developed turbulence in the near solar region, nonlinear processes are clearly central. It is conceivable that the wave models (based essentially on linear theory) could be completely misleading.

Interplanetary waves and turbulence are thought to play a significant role in governing a number of processes in the interplanetary medium. For example, even though the solar wind is quite collisionless, the  $\text{He}^{++}$  generally flows faster than the  $\text{H}^+$ , by as much as the local Alfvén speed. This fact is probably a signature of interaction with waves. At 0.3 AU the associated energy flux may be substantial. Thus, the helium may play an important role in the energy balance, which has not yet been elucidated. Another unexplained class of ion observations is that the ionic kinetic temperature tends to be higher than the proton kinetic temperature by a factor of the order of the ion mass ratio. This behavior is probably another manifestation of wave/turbulence.

Although there is convincing evidence for the solar origin of the Alfvénic component of the interplanetary fluctuations, other kinds of hydromagnetic fluctuations may be generated locally. A familiar example is the large-amplitude compressive turbulence associated with the interface between fast and slow solar wind streams. The firehose or mirror instabilities may regulate the anisotropy of the interplanetary plasma. There is ample evidence for hydromagnetic waves generated by particle beams from planetary bowshocks. The beams in turn are eventually disrupted by the waves they have generated. Hydromagnetic waves may be significant in the acceleration and scattering of energetic charged particles by Fermi acceleration and pitch-angle scattering. Waves and turbulence may also play a large role in governing transport processes involving ions and electrons of the main solar wind plasma. These and many related topics have been discussed over the past two decades, but very little has been accomplished in the way of verifying the operation of specific mechanisms.

## CHAPTER 5 - COUPLING OF THE SOLAR WIND TO THE MAGNETOSPHERE

A simple assessment of our present understanding of solar wind-magnetosphere coupling can be made in terms of the four major questions that were posed. The first of these questions is concerned with the process of solar wind plasma entry. Although solar wind plasma entry is not totally understood we do have observational evidence for the efficient injection of plasma through the cusps and into the plasma mantle. The open magnetic topology which is likely associated with reconnection and flux transfer events is at least partially responsible for cusp entry and possibly also for some plasma transport into the low-latitude boundary layer. Impulsive plasma injection also appears to be a quite reasonable hypothesis for the localized efficient injection of plasma into the low-latitude boundary layer. Finally, a number of cross-field diffusion processes have been proposed as closed model entry processes.

Further progress on plasma entry will require energetic particle and plasma measurements with high time resolution and high resolution in phase space, coupled with electric and magnetic field and plasma wave measurements with good coverage of the dayside magnetopause. Full 3-D plasma and particle distribution measurements are especially needed in the region of the outer cusp and the dayside boundary layer poleward of the cusp. Moreover, continuous measurements of the properties of the upstream solar wind in the vicinity of the Earth are critically important for future experimental studies of the solar wind-magnetosphere interaction. The different models that have been proposed to explain plasma entry must be refined into theories that can

be tested by plasma and field measurements. Since reconnection is a major candidate process, suitable modifications to the original Petschek (1964) theory should be made to treat the three dimensional aspects of the solar wind interactions and to examine the kinetic theory of the reconnection diffusion regions.

The second question concerns the processes of energy and momentum transfer from the solar wind to the magnetosphere. In this area we know that both open- and closed-model processes are important. The open models that result from steady-state reconnection and flux transfer events readily provide qualitative explanations for the response of magnetospheric convection and substorm activity to changes in solar wind parameters. On the other hand, closed-model processes are needed to explain the minimum polar cap potential drop of some 15 kV. This minimum potential is likely associated with the low-latitude boundary layer acting as an MHD generator on closed magnetic field lines. The suitability of the impulsive plasma injection model should be more fully explored and tested against spacecraft plasma measurements near the magnetopause. ISEE satellite measurements of waves and structure near the magnetopause provide valuable new information relevant to momentum transport via mechanisms such as the Kelvin-Helmoltz instability.

The third question deals with the physics of magnetospheric boundary layers, specifically their role as generators, loads and plasma transport regions. A major result of recent spacecraft programs has been the establishment of the existence and importance of these boundary layers. A more advanced level of understanding of magnetospheric boundary layers requires high-resolution measurements of their structure and dynamics. Theories of boundary-layer formation need to account for both MHD and kinetic effects and to establish the relationship of these effects to time-varying solar wind parameters.

The final question concerns the global magnetohydrodynamics that characterize the magnetosphere for the various coupling processes and as functions of solar wind parameters. Already the MHD models are capable of approximately describing the coupling between various magnetospheric regimes. As a complement to the global MHD models, magnetostatic equilibrium models are capable of describing the configuration of the quiet magnetosphere with higher spatial resolution. Both approaches to global modeling have now been extended to treat the system in three dimensions. Development of complete models will require an enhanced ability to specify coupling processes and boundary conditions such as the plasma distributions and current systems that characterize the various magnetospheric boundary layers.

The variety of magnetospheric phenomena discussed in this chapter leads to a number of unresolved physical questions that need to be answered in the light of observations, as well as from a theoretical point of view. In order to advance our present understanding, we recommend consideration of the following items:

1. There is observational evidence that the plasma near the magnetopause boundary, and near the plasma sheet boundary as well, is often highly anisotropic. Moreover, we know that some plasma drift mechanisms ultimately lead to an energy dispersion of the thermal plasma. Consequently, global MHD models for the magnetosphere, and magnetohydrostatic equilibrium models as well, should be formulated eventually in terms of multifluid codes and in terms of anisotropic thermal plasma pressure.
2. There is an overall agreement that the interaction between the plasma and the magnetic field within discontinuity regions (e.g., the magnetopause and the plasma sheet boundary layer) is dominated by microphysical plasma processes. The global fluid models need macroscopic physical quantities in their basic equations. It is, therefore, necessary to have a global model for

the resistivity at the magnetopause. Such a model could be obtained from observations or derived from the kinetic theory. For example, information is essential for a proper calculation of the normal magnetic field component  $B_n(x)$  at the magnetopause.

3. At the magnetopause, the quantity  $j(x) \cdot E(x)$  is a measure of the conversion rate of magnetic into mechanical energy and vice versa. The Chapman-Ferraro current density  $j(x)$  is determined uniquely by the distribution of the magnetic normal component  $B_n(x)$  at this boundary. Is it possible to derive macroscopic boundary values  $B_n(x)$  on the basis of kinetic arguments (see item 2) in order to calculate the global energy conversion rate at the magnetopause?
4. The magnetic normal component  $B_n(x)$  determines the "openness" of the magnetosphere, the size of the auroral oval, and the ionospheric convection  $E$ -field pattern. The polar cap electric field is assumed to reflect the solar wind  $(v \times B)$  field under the assumption that the ideal MHD theory is applicable. We know, however, that there exist significant magnetic field-aligned electric fields along the auroral oval. To what extent do field-aligned electric fields affect the polar cap convection  $E$ -field which is mapped from the solar wind through the magnetopause down to the ionosphere?
5. Sources for the plasma sheet are thought to include the ionosphere, the low-latitude boundary layer, and the plasma mantle. Are these known plasma sources sufficient to account for plasma sheet replenishment on substorm time scales?
6. Can we improve measurements and understanding of the minimum polar cap potential during extended quiet periods with northward interplanetary magnetic field (IMF)?
7. From the standpoint of global modeling as well as from our understanding of the time-dependent reconnection process, it is vital to obtain improved kinetic theory of the diffusion regions at the reconnection site.
8. Are flux transfer events the "debris" or active signatures of reconnection? Are they possibly associated with "impulsive injection events" or finite plasma filaments?
9. Can statistical analyses using solar wind and magnetospheric parameters be used to compare the relative efficacy of driven versus indirectly driven models of magnetospheric response and of closed versus open models of coupling? The results of such studies could act as a guide to the development of better theoretical models of solar wind-magnetospheric coupling.
10. Does the average plasma flow in the plasma sheet correspond to the predictions of MHD convection models? Observations of this flow and of  $B_z$  can give the pattern of the convection electric field. This pattern is likely to be quite different for open- and closed-coupling processes. Thus, these observations can provide a check on the importance of both processes as well as on the validity of convection models.
11. A tangential component of the electric field at the magnetopause is predicted as a consequence of magnetic field reconnection, and thus electric field measurements at the magnetopause are especially important. Special efforts should be made to develop and apply improved definitive measurements of vector electric fields at the magnetopause.

Our evaluation of the solar wind-magnetosphere coupling problem leads us to emphasize the importance of the boundary layers as the primary transport and coupling regions for the magnetosphere even though they constitute <5% by volume of the total system. We recommend that future observational and theoretical studies should focus on these boundary regions to determine the boundary conditions that they impose on the magnetospheric system and to isolate the dominant coupling processes.



## CHAPTER 6 - CORONAL TRANSIENTS AND THEIR INTERPLANETARY EFFECTS

A statistical association between solar flares and geomagnetic storms has been widely accepted since early in the 20th century, and the interpretation of this effect in terms of an expulsion of material from the Sun has been a commonplace in the field of solar-terrestrial physics for nearly as long. In situ observations of outward-propagating interplanetary shock waves made in the 1960s provided direct evidence for such a phenomenon. The masses inferred from the shock wave observations implied that their passage through the solar corona should produce major perturbations of that region. Thus the first identifications of coronal transients or mass ejections by spacecraft-borne instruments in the early 1970s were of interest to a far broader group of scientists than those who specialized in this most tenuous domain of solar physics. The role of coronal mass ejections in solar-terrestrial physics, or more specifically, in the chain of cause and effect introduced above (solar flare-interplanetary shock wave-geomagnetic storm), has remained an important, continuing theme in the study of mass ejections. Analysis of the "first generation" of these coronal observations, development of theoretical models of mass ejections, and comparisons with both solar and interplanetary data have yet to provide definitive answers to the first two questions that arise in pursuit of this theme: (1) What are the solar origins (in both a phenomenological and physical sense) of coronal mass ejections, and (2) What are the interplanetary effects of coronal mass ejections?

These two fundamental and familiar questions concerning the role of coronal mass ejections in the context of solar-terrestrial physics were the foci of our discussions at this workshop. In this document, we first describe some of the necessary background material on flares and geomagnetic storms, on interplanetary shock waves, and on coronal mass ejections themselves. We then describe one of the modern "tools" available for approaching these questions -- theoretical models for the initiation and propagation of transient phenomenon in the solar corona. All of this material has been extensively reviewed in the recent literature, and our coverage of it will be both selective and somewhat abbreviated. We then describe the second of our "tools" -- a new generation of coronagraph observations of mass ejections and complementary set of solar and interplanetary observations suitable for correlative studies.

Several studies of solar wind data have the potential of both improving our understanding of the interplanetary effects of coronal mass ejections and yielding information on the physical mechanisms of the ejections themselves. Some of these topics can be pursued using existing data bases but others require development of new experimental opportunities using state-of-the-art instrumentation.

Magnetic Topology of Driver Gas. Coronal images of mass ejection events show nested loops of bright material moving away from the Sun. At present it is not clear whether these loops represent density-loaded magnetic field lines or shock fronts. If they outline magnetic field lines, as is often assumed, the role of magnetic reconnection before, during and after "lift off" reappears as an important question. The bearing of this question on the mechanism driving the mass ejection suggests considerable attention in future research. Quantitative analyses of existing solar wind and energetic ion data may provide some useful clues in this regard.

Heavy-Ion Spectroscopy. The ionization state of the solar wind provides information about electron temperatures in the corona between about 1.1 and ~6 solar radii. Any such information bears on the processes that expel coronal plasma. Interpretation of these data would be facilitated by analysis of simultaneously observed soft X-ray images of coronal regions in the neighborhood of disappearing filaments.

Theoretical Models of Chemical Fractionation Mechanisms. A characteristic feature of solar wind plasma associated with the shock-wave driver gas and hence some coronal mass ejections is enhanced yet variable  $\text{He}^4$  abundance. Although detailed analyses have not yet been completed, a large data base of  $\text{He}^3$  abundances measured aboard ISEE-3 is available to investigate whether  $\text{He}^3$  varies in a manner similar to  $\text{He}^4$ . Regardless of the outcome of such a study, the causes of the known, strong  $\text{He}^4$  variations must bear on conditions in, and/or dynamical processes operating within the solar atmosphere near the sites of coronal mass ejections. However, very little is presently known about the mechanisms which can cause chemical fractionation in such regions. This is a wide open field of theoretical endeavor and any information could help organize a large body of data that may prove relevant to the physical mechanisms leading to mass ejections.

Correlative Studies of Coronal Ejections and Their Interplanetary Effects. A fundamental uncertainty in the study of coronal mass ejections is their three-dimensional geometry. Neither 2-D measurements such as provided by coronagraph and X-ray images nor 1-D measurements such as provided by in situ solar wind measurements have resolved this problem. However, a combination of pertinent coronal white-light and X-ray images, type II radio emission intensities, satellite-source radio scintillation measurements, and multipoint in situ measurements are available. A coordinated analysis of these data could yield much detailed information concerning the nature and spatial structure of the driver gas, the spatial relationship between the bright loops so prominent in coronagraph images and a possible shock front near the corona, the 3-D global extent of the resulting interplanetary disturbances and the effectiveness of coronal mass ejections in generating interplanetary shocks.

Type II Generation Mechanism and Spatial Extent. Before type II radio emissions can be used as a sensitive quantitative tool for exploring coronal mass ejection mechanisms and their interplanetary effects, it is first necessary to understand the physical mechanisms which convert a plasma disturbance to radio waves. Although there is overwhelming evidence to indicate that type II emissions come from shock wave disturbances, the relationship known from in situ measurements is not one-to-one. However, recent advances in our understanding of the processes which lead to electron acceleration and wave generation in collisionless shocks lead us to believe that a concerted effort to understand type II emission mechanisms using ISEE-3 data should lead to substantial progress.

## CHAPTER 7 - CONNECTION BETWEEN THE MAGNETOSPHERE AND IONOSPHERE

Two decades of space research have produced ample evidence that particles and fields originating in the active Sun can gain entry into the terrestrial magnetosphere and deposit their energy in the ionosphere and atmosphere. The final link in this solar-terrestrial chain is generically referred to as *magnetosphere-ionosphere coupling* (MIC). Included in the physics of the MIC processes are not only plasma processes associated with the earthward deposition of energy of the active Sun but also the response of the ionosphere-atmosphere system to the energy deposition. Because of its key position in the physics of the near-Earth space environment, MIC processes have been a major focus of space research.

Up until the past several years, the important issue in MIC was the role of the ionosphere in magnetospheric field configuration of the global scale. Through comprehensive formulations of MIC theory the schematic relationships between magnetospheric magnetic and electric configurations and ionospheric plasma flow have been elucidated. Thus, the ionosphere-atmosphere, through dissipative processes inherent in the finite ionospheric conductivity, is thought to regulate the electrodynamic circuit driven by the interaction between the magnetosphere and the solar wind at the magnetospheric boundary.

Despite the highly successful elucidation of the schematic features of MIC by global-scale theory, it soon became clear that MIC involves much more complex plasma kinetic effects. This significant change of theoretical perspective since the latter part of the 1970s was due to plasma and field observations by auroral satellites in the  $\sim 1 R_E$  altitude region and by rocket chemical releases. New observations from these auroral satellites have not only confirmed previous suspicions that a parallel electric potential drop may interface with global MIC processes but have also introduced entirely new kinetic phenomena such as upflowing ionospheric ions and complex wave-particle interactions. These are likely to change our perspective of the role of the topside ionosphere in MIC processes. Recent observations of polar region auroral phenomena have indicated that our picture of magnetospheric topology may need to be revised. In short, we have witnessed in the last 4-5 years a dramatic change in perspective on the physics of magnetosphere-ionosphere coupling.

Because of the far-reaching implications of recent discoveries in MIC and because a coherent assessment of them has yet to be made, the principal task in this group was to critically assess these new observations and the new perspectives that they may engender.

Based on our assessment of the progress in magnetosphere-ionosphere coupling physics in the last five-six years, we envisage a period in the near future in which basic and global issues in MIC will be addressed observationally and theoretically. These expectations are derived from the following general conclusions:

- Magnetosphere-ionosphere coupling (MIC) encompasses a complex set of interacting phenomena occurring on vastly different time and spatial scales rather than a set of isolated processes.
- Energy flow in the ionosphere-magnetosphere system is primarily from the magnetosphere to the ionosphere. However, it is now clear that the ionosphere is not a passive element in this system. The spatial distribution of ionospheric conductivity affects magnetospheric plasma transport and the configuration of the magnetospheric current systems. Atmospheric motions may significantly modify magnetospheric plasma circulation. Furthermore, a substantial portion of the plasma population of the magnetosphere originates in the ionosphere.
- Theory has developed considerably along with observation for electrodynamic coupling and related turbulent phenomena. However, the present state of theoretical development does not adequately address the feedback onto the magnetosphere from the dynamic ionospheric response to these processes. Currently, such theoretical analyses are just beginning but are required for complete understanding of magnetosphere-ionosphere coupling.
- The global scale theories of magnetosphere-ionosphere coupling have matured to the point that it is possible to test their assumptions and predictions. Thus, experiments can be designed to test specific aspects of our theoretical understanding of global scale magnetosphere-ionosphere coupling.

While the above general conclusions are concerned with the scheme of magnetosphere-ionosphere coupling as an integrated whole, we have also identified the following pressing issues which address specific areas of MIC:

- Global imagery of the aurora in optical, UV and X-ray is needed to determine a coherent picture of the global configurations of MIC.

- What is the specific spatial and temporal relationship between global-scale field-aligned current ( $J_{\parallel}$ ) systems, auroral potential structures, and auroral arc structures?
- What and where are the generators of the entire scale spectrum of field-aligned currents? In particular, is the MIC dynamo a voltage generator or a current generator? What is the nature of the auroral return current? How does upper atmospheric circulation affect magnetospheric current flow?
- Because auroral phenomena are generally associated with topological boundaries of magnetospheric regions, three-dimensional models of the magnetospheric electric field structure is needed.
- Quantitative models of ionospheric influence upon magnetospheric phenomena are needed. In particular, the importance of the ionospheric plasma source and of the influence of ionospheric conductivity upon MIC dynamics must be addressed quantitatively.
- At present there is a paucity of theories of auroral morphology in relation to MIC processes in the dayside and in the post-midnight sectors.
- It is by now clear that auroral plasma kinetic characteristics show both adiabatic and diffusive signatures. However, the feedback effects of wave turbulence generated by the auroral electron beam upon the distribution of  $E_{\parallel}$  have not been quantified.
- Currently, there is a noticeable lack of theories of dynamic auroral processes such as the westward traveling surge.
- Since auroral arcs at very high latitude regions can involve plasmas from the magnetosheath and from the plasma sheet, quantitative theories of formation for such arcs are needed to understand their relationship to the oval arcs.
- Auroral kilometric radiation (AKR) reflects detailed structures of plasma processes of its source region, therefore, AKR can be used as a remote diagnostic tool of MIC processes. Once terrestrial AKR processes are understood, this diagnostic tool can be applied to MIC processes in other planets.
- Auroral wave-particle interactions depend on kinetic characteristics of the plasma. Some of these characteristics may result from the plasma transport itself, therefore, it is important to quantitatively understand the kinetic theory of auroral plasma transport.

## CHAPTER 8 - SUBSTORMS IN THE MAGNETOSPHERE

In this report a working definition of substorms was developed, the primary relationships within the solar wind-magnetosphere-ionosphere system as they relate to substorms were considered and the role that substorm studies play in the overall discipline of solar-terrestrial research was clarified. In pursuing the last of these topics, the working group addressed the question of which aspects of magnetospheric activity are directly driven by the solar wind and which aspects of this activity represent an unloading process for previously stored solar wind energy.

Substorm studies really comprise a global examination of the interrelationships between many detailed physical processes (see the reports of Working Groups 1, 2, 5, 7, and 11). Furthermore, substorms are a complex combination of the driven and unloading processes which are described in some detail in the body of this report.

Significant progress has been made in understanding substorm phenomenology and, to a large extent, in understanding the underlying physical mechanisms. Nonetheless, considerable disagreement and uncertainty remains. In studying substorms, we are struggling with a truly staggering global problem wherein experimental probing and/or theoretical modeling must deal with extreme variations in scale sizes. To a very large degree, the era in which a single scientific instrument on a single spacecraft can add substantially to the knowledge of substorm processes is past. Global measurements, combined with realistic global 3-D numerical simulations, probably hold the greatest promise of advancing our understanding significantly.

Substorm research has established that a variety of phenomena (e.g., auroral break-up, injections at synchronous orbit, and reconnection signatures in the distant magnetotail) occur at widely separated locations in association with the expansion phase onset. Further research -- both theoretical and observational -- is needed to gain a more detailed, quantitative understanding of the interrelationships among these various phenomena. In particular, the following questions need to be addressed.

- What is the nature of the energy transfer mechanisms between the presumed energy storage reservoir in the magnetotail and the various sites of energy dissipation? Proposed transfer mechanisms such as compression waves and field-aligned currents have been developed only qualitatively to date.
- How do these energy transfer mechanisms relate to and/or arise out of the reconnection process?
- What are the propagation time delays between the magnetotail and the inner magnetosphere/ionosphere, and what ramifications do these delays have for substorm dynamics?

A complete understanding of any given magnetospheric process requires that the role of that process in the global dynamics of the magnetosphere be identified. An essential aid to attaining this global perspective is the world-wide array of all-sky camera, magnetometers, riometers, and other monitors of magnetospheric processes. Note that all points in the outer magnetosphere -- roughly, beyond synchronous orbit -- have their magnetic footprints in auroral and polar regions. A suitably spaced array of stations in these regions thus provides a vital means of continuously monitoring the state of the magnetosphere on a global scale. To fully exploit this monitoring capability, further progress is needed in identifying the ground signatures of specific magnetospheric processes and in improving ground-based indices of energy flow and dissipation.

The auroral electrojet indices (AE, AU, AL) have been extensively used as measures of the auroral electrojet intensity and magnetospheric substorm activity. Their accuracy and limitations have been discussed to some extent by several authors but there had been no systematic way to evaluate or calibrate them. The AE index is an important and useful index when it is used properly, but one must be carefully aware of, and take into account, its limitations. During recent years, the AE(12) index has been used frequently for purposes which require accuracy well above that of the presently available index.

We firmly believe that future correlation studies between the AE index and solar wind quantities must be based on an improved AE index. Moreover it is crucial to continually improve geomagnetic indices and estimates of the global energy dissipation rates for there to be new advances in this particular area of magnetospheric substorm studies.

One of the most important problems in substorms studies has been to examine how the energy accumulated in the magnetotail varies during magnetospheric substorms. Satellite measurements of the magnetic field magnitude,  $B$ , provide a measure of the magnetic energy density ( $B^2/8\pi$ ) which may be characteristic of the entire volume of the magnetotail. However,  $B$  is sensitive to changes of the solar wind pressure and of the local plasma pressure. Furthermore such data usually represent only a single point measurement and the data are only available for limited periods. Therefore, it is important to look for a measure of tail magnetic energy, which is continuous and which is not subject to great uncertainty. One such new measure may be the "diameter" of the auroral oval, since the total open flux in the tail is expected to be roughly proportional to  $B_G(d/2)^2$  (where  $B_G$  denotes the vertical component of the magnetic field in the polar cap region) and since an increase of the diameter  $d$  indicates an increase of the open flux.

In a related vein, a phenomenon that is observed with striking regularity in the magnetotail beyond  $r \approx 15 R_E$  is the thickening or recovery of the plasma sheet during the subsidence of negative bays at nightside auroral zone stations. In fact, the plasma sheet thickening sometimes seems to occur in very close coincidence with peak bay activity and with the beginning of final recovery of auroral zone bays and with the onset of bays at low polar cap locations. Thus it has been inferred that a "poleward leap" of the electrojet occurs in coincidence with the plasma sheet's recovery is ascribed to a sudden tailward retreat of the substorm neutral line to large distance. Lobe plasma, carried by convection into the neighborhood of the neutral line and jetting earthward from it, threaded with contracting, newly reconnected field lines is thought to be the material reconstituting the plasma sheet. Theoretical and modeling studies should be directed toward understanding the retreating neutral line phenomenon.

The conventional description of the auroral substorm contains no reference to the poleward leap, but has the auroras moving to high latitudes during the expansion phase and then retreating gradually to lower latitudes during a "recovery" phase. The later phase of substorms needs to be studied with the advanced equipment now available (magnetometer chains, auroral imaging from satellites, high-latitude auroral radars) to determine more precisely the features that are attendant upon the retreat of the neutral line and plasma sheet recovery.

The interface between the plasma sheet and the tail lobe is a region where very dynamic processes are often found including plasma jetting, beams of energetic particles, and intense plasma waves. These processes are particularly prominent as the interface (i.e., the boundary of the plasma sheet) surges over a satellite during plasma sheet recovery. It is not surprising, in light of the neutral line model of substorms, that this interface should be a site of activity since it contains magnetic field lines connected directly to an X-type neutral line. This interface should receive detailed study both to derive information it may reveal regarding the reconnection process and to further our understanding of the substorm sequence, particularly the phenomena of plasmoid ejection and the retreating substorm neutral line.

As we consider substorm energy dissipation, we realize that a reasonable first-order estimate can regularly be made for ring current injection rates, Joule heating rates, and auroral particle precipitation rates. Since these dissipation rates can be (at least roughly) estimated, they can be related to quantitative energy input estimates derived from dayside reconnection models, viscous interaction models, etc. A major remaining deficiency as we consider the partitioning of the available input energy is determining (on a regular and routine basis) how much energy simply escapes down the tail in the plasmoid structure. Methods should be developed to monitor this component of substorm energy dissipation.

In the area of injection and energization of hot plasma in the outer equatorial magnetosphere, several problems remain outstanding. Among these are:

- What is the role of global convection relative to the role of impulsive injections in the formation of the ring current?
- What is the nature of the earthward-propagating compressional waves observed in association with impulsive injection events? In particular, can present models of tail collapse give rise to such a wave?
- Are there other mechanisms that can be invoked, beside the observed compressional wave, to account for impulsive, dispersionless injection events?
- What is the relationship between the injected hot equatorial plasmas and the field-aligned currents which occur in association with substorm auroral displays?

To attack these and other questions in this area, we must experimentally determine the detailed spatial structures of the injection boundary/injection front. Furthermore, very detailed multiple-satellite timings of events in the tail, in the near-geostationary environment, and on the ground must be made in order to relate the earthward injection of hot plasma to substorm processes in the more distant tail. Detailed numerical models being developed to describe tail dynamical processes should be modified to include compressional wave propagation (if possible) and hot plasma injection characteristics. Finally, detailed statistical and event analyses need to be performed in order to relate hot plasma distributions to field-aligned current structures and ring current development.

Numerous other problems remain to be addressed. For example, the mechanisms for triggering plasma sheet instabilities that may be responsible for substorm expansion onsets remain largely undetermined. Detailed theoretical analyses of the possible role of ionospheric ions in lowering plasma sheet instability thresholds should be undertaken and appropriate experimental investigations should be performed. In fact, the entire area of heavy ions in magnetospheric dynamics is just opening up and, furthermore, the appreciation of positive feedback loops (as those involving the ionosphere) is only beginning to come about. As observations have improved and as theory has progressed there has developed a further realization of the very significant analogy which exists between the explosive acceleration of energetic particles in the Earth's magnetosphere during substorm onsets and on the Sun during solar flare onsets. As we compare and contrast various parts of the solar-terrestrial system, we should more fully exploit such analogous processes in sub-disciplines that are often far-removed from one another.

## CHAPTER 9 - IMPACT OF FLARES ON THE SOLAR TERRESTRIAL ENVIRONMENT

The solar flare is one of the most dramatic phenomena in nature. A large amount of energy, in the form of photons and plasma and energetic particles, is released suddenly from the solar atmosphere and impacts the entire solar-terrestrial environment, from the Sun, through the heliosphere to the Earth and beyond. Flares generate natural phenomena such as shock waves and geomagnetic storms; flares affect man's technology and his ability to survive in space. It is the purpose of this chapter to assess our current understanding of these impacts.

A study of the impacts of solar flares encompasses many of the phenomena that are considered in detail in other chapters of this document. Solar flares are examples of particle acceleration, solar flares affect solar wind-magnetosphere coupling and magnetosphere-ionosphere coupling. We are concerned here only with the effects which are the direct and unique results of solar flares.

Solar flares impact the entire solar-terrestrial environment and cause a host of phenomena, some of which have received considerable study in recent years. Yet despite this study, there are some fundamental points, and many missing details of the impact of solar flares, that remain to be understood which include:

- The flare process yields electromagnetic radiation at all frequencies from radio to gamma rays, energetic particles from solar wind energies to GeV, and enhanced solar wind flow. This radiation and particles provide us with detailed diagnostics of the flare process. Yet with the current observational techniques we do not know the exact location where the flare energy is released, and as a result the exact physical conditions at the flare site. More detailed, higher resolution observations are necessary. These observations can serve as the basis for comprehensive theories of the energy release in solar flares and for shock generation, which do not currently exist, and ultimately for developing models to predict the occurrence of solar flares.
- Solar flares impact the heliosphere primarily through the enhanced solar wind flow that they generate, and the accompanying shock waves. These flows and shocks can heat the solar wind, modulate and accelerate energetic particles, and cause deviations in the direction of the interplanetary magnetic field out of the ecliptic plane which cause geomagnetic disturbances. More comprehensive theories need to be developed to describe the dissipation of enhanced solar wind flows into heat, and the interaction of energetic particles with these flows. For example, theories of galactic cosmic ray interaction are very rudimentary; theories for shock acceleration of energetic particles have developed rapidly in the last several years but need now to be applied in comprehensive models of realistic shocks. Similarly, the mechanisms for causing variations in the interplanetary magnetic field that can drive magnetospheric phenomena need to be understood more fully. Such understanding is an essential component in predicting the impact on the magnetosphere of a given flare.
- Solar flares are responsible for a wide variety of magnetospheric phenomena: they can generate storms and substorms, supply energetic particles to the magnetosphere, and influence the adiabatic and non-adiabatic motion of trapped particles. For some of these phenomena there are conceptual uncertainties that must be resolved; for example, the microscopic plasma processes that are responsible for solar wind energy conversion to magnetospheric phenomena need to be understood. Other phenomena are conceptually understood, but not in detail, e.g., aspects of storm and substorm dynamics, and details of charged particle transport and energization require study.
- Solar flare EUV and X-rays affect the ionospheric electron content and D-region chemistry; solar flare particles affect D-region ionization, ion/neutral chemistry including ozone, and atmospheric electricity. The physics of most of these effects is understood, but their global implications remain to be adequately modeled. For example, an adequate assessment and treatment of global sources, such as magnetospheric electron precipitation, or of global transport of disturbances, needs to be developed. Global modeling is essential if we are ever to determine whether solar flares, or any transient solar effect, influences the weather and/or the climate.
- Solar flares affect many activities in our technological world: radio, satellite and cable communication, power transmission, pipeline corrosion, and manned and unmanned space-flight. Needed research in this aspect of solar flare impact is difficult to predict, other than to note that as technology becomes more sophisticated our knowledge of the natural environment which can impact that technology must become more detailed. For example, we must improve our modeling of the radiation belt environment of the magnetosphere to



satisfy the current and expected needs of spacecraft designers; we must determine the scale-sizes of transient ionospheric and magnetospheric currents to set the design specifications for long, groundbased conductors.

## CHAPTER 10 - COLLISIONLESS SHOCK WAVES IN THE SOLAR TERRESTRIAL ENVIRONMENT

Shock waves are created by the nonlinear steepening of compressive wavemodes in a fluid. Typically such steepening occurs when a disturbance travels through the fluid at a speed higher than the characteristic speed with which small amplitude compressive waves can propagate. The shock formed in front of such a disturbance is the means by which the disturbance communicates with the ambient plasma into which it is propagating. At the shock the flow normal to the shock must be changed from super-"sonic" upstream to sub-"sonic" downstream, where "sonic" refers to the characteristic speed of the small amplitude pressure signals. In plasma this speed is usually the magnetoacoustic speed. The loss of streaming energy represented by this slowing of the flow at the shock is converted into other forms of energy. Broadly speaking a shock may be defined as the entire region over which any portion of the conversion takes place. This conversion must be accomplished via collisionless dissipation mechanisms. Identifying and understanding dissipation mechanisms is therefore a central control of many shock studies.

Laboratory experiments, numerical simulations, and space observations all indicate that the modes of dissipation at a shock depend sensitively upon upstream flow conditions, particularly the Mach number of the flow,  $M$ , and the angle between the local shock normal and the upstream magnetic field,  $\theta_{Bn}$ , and the ratio of thermal to magnetic internal energy,  $\beta$ . Shocks with Mach numbers below some critical value,  $M_C$ , are known as subcritical shocks, whereas those with  $M > M_C$  are known as supercritical.  $M_C$  itself depends upon such things as the upstream plasma and  $\theta_{Bn}$ , but is usually within the range 2-3. The fundamental difference between these two classes of shocks is that dispersion and anomalous resistivity provides most of the dissipation for subcritical shocks whereas ion reflection at the shock plays an increasingly important role in the dissipation process as the Mach number increases above  $M_C$ . A fundamental change in shock structure occurs also at  $\theta_{Bn} \sim 45^\circ$  which is most evident as an increase in turbulence and the lack of a well defined shock structure when  $\theta_{Bn}$  is  $\leq 45^\circ$ . This provides a rough dividing line between two further classes of shocks: quasi-perpendicular ( $\theta_{Bn} > 45^\circ$ ) and quasi-parallel ( $\theta_{Bn} < 45^\circ$ ).

Small amplitude waves can propagate through a plasma with different characteristic speeds depending upon the mode of propagation. This difference leads to a further classification of shock structure - slow, intermediate, and fast which, correspond to disturbances exceeding respectively the slow magnetoacoustic speed, the Alfvén speed, and the fast magnetoacoustic speed. Only fast shocks have been studied extensively; however, slow shocks have been observed in the solar wind and may play an important role in field line reconnection processes. In this report we concentrate exclusively on fast shocks, unless stated otherwise.

Collisionless shocks commonly contain a structural element known as the foreshock which is not present in ordinary collisional shocks. The foreshock extends for a considerable distance upstream from the main shock transition and contains suprathermal ions and electrons reflected at the shock or transmitted through it from the downstream region. These ions and electrons excite a variety of hydromagnetic and plasma waves, which in turn scatter the particles. As the wave speeds are generally lower than the convection speed of the upstream plasma flow, the waves are subsequently convected back toward the shock. Ultimately, scattering of the initial "seed" populations of reflected and/or leaked particles off of these waves both up and downstream from the shock can produce a significant acceleration of a small fraction of the ions to very high energy. Such acceleration appears to be a ubiquitous feature of all shock waves observed within the heliosphere.

Progress in our knowledge of collisionless shocks, with particular attention to the important contributions made by satellite instruments, has been reviewed extensively in the literature. In this report, we present a summary of selected information pertaining to the current status of our understanding of the physical processes in shocks. Much has been learned lately; even more is being studied and prepared for publication as this is written. Exposing the newest results and their relationships to each other is a major aim of this report.

The working group developed the following list to summarize the foremost outstanding problems, more or less in decreasing order of priority, that are expected to engage investigators of collisionless shock physics in the foreseeable future.

#### Quasi-parallel Shocks

What is their full field and particle structure?

What role does leakage of downstream into upstream plasma play in their profile?

How are their intrinsic properties separated from those dependent on communication with other sections of the same shock?

#### Quasi-perpendicular, turbulent shocks

How (by what instabilities) are ions actually thermalized initially in the shock ramp and more completely behind it?

Is there an electrostatic subshock?

Is there a second, supercritical Mach number? If so, what is it, what does it depend on, and how does it affect shock structure?

Where are the sources of seed ions and electrons for the foreshock?

What role do multiple reflections of ions play in the reflection process?

#### Quasi-perpendicular, laminar shocks

How are ions heated in the laminar shock?

Are standing waves stable? If not, what becomes of them?

#### Parametric Structure

What is the full, definable, macroscopic, quantitative categorization of shock structures by  $\beta$ , Mach number, etc.? What nomenclature should be standardized to describe them, e.g., quasilaminar; quasi-turbulent?

### Ion Acceleration

How, if at all, do suprathermal ions at interplanetary shocks differ from those at the bow-shock?

### Electrons

What is the detailed origin of the flat-topped distributions?

What control does local geometry have on potential-drop heating of electrons?

### Waves

What is the origin of ion sound turbulence?

What role do nonlinear wave-wave interactions play in shock wave dissipation?

How, if at all, do nonlinear plasma waves participate in the thermalization process?

### Global Aspects

What is the large scale closure system of the shock currents?

Can (should?) structural factors be incorporated in the jump conditions in multifluid, MHD flow codes for planetary shocks?

To what extent does quasi-parallel/quasi-perpendicular geometry and its variability affect the overall solar wind-magnetosphere interaction?

The study of collisionless shocks can, and has been, pursued by laboratory experimentation, theory/simulation, and natural observation. Recent years have seen an abundance of observational data, and a rising use of numerical simulations. In contrast, laboratory experimentation has been abandoned, and purely analytic theory has been substantially curtailed. It is clear that in the immediate future the predominance of observation and simulation will continue. Nevertheless, each of these approaches has its advantages and its limitations, and none should be excluded from future consideration.

## CHAPTER 11 - ASSESSMENT OF PLASMA TRANSPORT AND CONVECTION AT HIGH LATITUDES

The high-latitude ionosphere is strongly coupled to the thermosphere and magnetosphere. The magnetospheric coupling occurs via electric fields, field-aligned currents, and particle precipitation. Owing to the interaction of the shocked solar wind with the geomagnetic field, an electric potential difference is generated across the tail of the magnetosphere, with the resulting electric field pointing from dawn to dusk. Except for isolated regions, typically in the auroral oval, the geomagnetic field lines are equipotentials due to the high electrical conductivity along field lines. Consequently, this cross-tail potential difference is mapped into the high-latitude ionosphere as an electric field that is directed perpendicular to the geomagnetic field. At ionospheric heights, this perpendicular (or convection) electric field is typically  $25\text{--}50\text{ mV m}^{-1}$  in the polar cap, but can be much greater than  $100\text{ mV m}^{-1}$  in restricted latitudinal bands at certain times.

The effect that the convection electric field has on the ionosphere depends on altitude. At F-region altitudes and above, both ions and electrons drift in the  $\underline{E} \times \underline{B}$  direction, where  $\underline{E}$  is the convection electric field and  $\underline{B}$  is the geomagnetic field. The convection pattern is basically a two-cell pattern with antisunward flow over the polar cap and return flow at lower latitudes. At low altitudes (E region), the ion velocity vector rotates toward the  $\underline{E}$  direction owing to the increased ion-neutral collision frequency, while the electrons continue to drift in the  $\underline{E} \times \underline{B}$  direction. Consequently, large-scale horizontal currents ( $J_P$  and  $J_H$ ) flow in the E region. These horizontal ionospheric currents are coupled to the magnetosphere through field-aligned (Birkeland) currents. The magnitude and direction of the currents as well as the heating associated with current dissipation depend not only on the magnetospheric generator, but also on the neutral gas motion. Likewise, the neutral gas is strongly affected by the convecting plasma through direct momentum forcing and frictional heating, which act to induce changes in the thermospheric wind and composition. These changes, in turn, affect the ionospheric densities, conductivities and currents.

Energetic particle precipitation from the magnetosphere in the auroral region leads to the creation of ionization and to electron, ion, and neutral gas heating. These processes, in turn, affect the ionospheric conductivities and currents as well as the neutral wind and composition. The precipitating electrons gain energy from electric fields that are directed parallel to the geomagnetic field, which occur only in the auroral region. The parallel electric fields also produce energetic, upstreaming  $O^+$  ions ( $\sim keV$ ), which are an important ionospheric source for the magnetosphere.

The ionosphere also is a source of both thermal ( $\sim 0.5$  eV) and suprathermal ( $\sim 5$  eV) ions via the polar wind. At altitudes greater than about 800 km, the ionospheric ions can escape from the topside ionosphere along geomagnetic field lines owing to the pressure difference ( $dp/dz$ ) between the ionosphere and magnetosphere; a process termed the polar wind. Substantial fluxes of  $H^+$ ,  $He^+$  and  $O^+$  ions continually flow into the magnetosphere on geomagnetic field lines greater than about  $60^\circ$  (invariant latitude).

In order to assess our current understanding of plasma transport and convection at high latitudes, it was necessary to take account of the strong coupling between the ionosphere, thermosphere, and magnetosphere.

To date, the different elements of the ionosphere-thermosphere system have generally been studied separately. That is, there have been many studies dealing with high-latitude currents and conductivities, thermospheric winds, plasma convection, and the polar wind, but the time-dependent coupling between the different elements has largely been ignored. Therefore, in the future, it is important to study the coupled system for a range of time scales, covering wave phenomena, diurnal, seasonal, and solar cycle variations.

In the course of studying the different elements separately, a number of major unresolved problems have surfaced. However, many of these problems may simply be related to the lack of consideration of the coupled, time-dependent behavior of the ionosphere-thermosphere system. At the present time, some of the major unresolved problems are the following:

- What is the relationship between the Birkeland currents, horizontal currents, and the convection electric field?
- What is the relationship between ionospheric convection and thermospheric winds? Does the wind drag or drive the ionosphere?
- Can the atmosphere-ionosphere dynamo affect magnetospheric transport?

- Can atmospheric gravity waves and magnetospheric MHD waves affect the ionospheric energy budget?
- How do the potential distributions and plasma convection patterns vary both spatially and temporally?
- What is the response of the global ionization distribution to plasma convection?
- What is the relationship of convection patterns in the conjugate ionospheres and neutral wind systems?
- Is the polar wind supersonic or subsonic?
- Is the polar wind stable?
- What fraction of magnetospheric ions is related to the polar wind?

## CHAPTER 12 - HIGH LATITUDE IONOSPHERE STRUCTURE

The Earth's ionosphere is an important element in solar-terrestrial energy transfer processes. As a major terrestrial sink for many solar and magnetospheric events, the ionosphere has characteristic features that can be traced to such seemingly remote phenomena as solar flares, radiation belt wave-particle interactions and magnetospheric substorms.

In considering the multitude of solar-terrestrial plasma interactions, it is important to recognize that the high-latitude ionosphere is not altogether a simple receptor of various energy deposition processes. The high-latitude ionosphere can play an active feedback role by controlling the conductivity at the base of far-reaching magnetic field lines and by providing a plasma source for the magnetosphere. Indeed, the role of the ionosphere during magnetospheric substorms is emerging as a topic for meaningful study in the overall picture of magnetospheric-ionospheric coupling.

The accessibility of the ionosphere provides the opportunity for rather detailed investigations with combinations of in situ diagnostics and remote sensing techniques. Thus it provides the plasma physics community exciting challenges in studying many of the dynamic and unstable ionospheric plasma states. Indeed, there are regions of the Earth's ionosphere that are more irregular than not, more dynamic than quiescent, and more unstable than previously appreciated.

The study of ionospheric structure, while traditionally focused on density irregularities, more generally includes irregularities in density, temperature, ion composition and fields. The study of irregularities in the ionospheric F region has divided itself naturally into three regimes: the low, middle and high geomagnetic latitudes. These regimes have fundamentally different sources for the irregularities, associated with the degree of coupling to higher altitude magnetospheric phenomena. The coupling is related, of course, to the orientation of the geomagnetic field as function of magnetic latitude. At low latitudes, where the geomagnetic field tends to be horizontal, coupling to higher altitude magnetospheric disturbances is inhibited. At high latitudes the more vertical magnetic field promotes strong magnetospheric-ionospheric coupling.

The term "ionospheric irregularities" has become synonymous with the terms "fluctuations", "structure", and "striations". Under varying conditions, ionospheric irregularities can populate a broad range of latitudes, longitudes and altitudes. The irregularities are generally magnetic field

aligned, i.e., there is little variation along the geomagnetic field. Thus, the irregularity wavelength perpendicular to the geomagnetic field is much smaller than its parallel wavelength and this two-dimensional nature of the irregularities simplifies most theoretical analyses. There can be a diversity of irregularity scale sizes, spanning up to six orders of magnitude (tens of centimeters to hundreds of kilometers). The existence of such a broad spectrum of irregularities is further complicated by varying geophysical conditions and wavelength-dependent mechanisms. While our understanding at high latitudes is still primitive, the past few years have seen exciting advances in our understanding of low-latitude ionospheric irregularities in the areas of equatorial spread-F and the equatorial electrojet. This advance has resulted from combined improvements in experimental, theoretical, and computer simulation techniques. The rapid development and meshing of these research tools in the past few years has made this a propitious time for unraveling some of the basic causal mechanisms of high latitude ionospheric irregularities.

The discussions of the working group identified several theoretical and experimental research thrusts which should be pursued to resolve some of the key questions. These major efforts are summarized in the following paragraphs.

### THEORETICAL THRUSTS

Plasma instabilities have been shown to play an important role in the generation and evolution of high-latitude ionospheric structure. Because plasma macroinstabilities become highly nonlinear, further quantitative progress in this area can only be made through numerical modeling and simulation of the fundamental plasma equations. Simple analytic nonlinear models would also be desirable not only to give insight into the basic nonlinear physics, but also to validate the output from large numerical simulations. Since the drivers of these plasma macroinstabilities are density gradients, electric fields, neutral winds, and field-aligned currents, it is important to quantify the scale sizes of the drivers themselves. In addition, since the high latitude F-region ionosphere is threaded by magnetic field lines, one must investigate the consequences of coupling of the F region to the E region and to the topside ionosphere and magnetosphere. Moreover, since the convective motions can be spatially dependent, it is important to assess the influence of velocity shear on the macroinstabilities. This includes modeling shears in the nonlinear numerical simulations. The possible seeding role of neutral field gravity waves in producing large scale irregularities deserves further investigation.

The generation of irregularities with wavelengths of centimeters to tens of meters in the auroral F region has been given considerably less theoretical attention. Several free energy sources may drive these microinstabilities. The basic properties of linear ionospheric instabilities due to density gradients and cross-field electric fields are now reasonably well established. The role of field-aligned currents is still being assessed. The effects of sheared cross-field flows and a sheared magnetic field on microinstabilities in the ionosphere are not understood and deserve substantial attention.

An outstanding question is the saturated amplitudes and spectra of the small-scale irregularities and the role they play in influencing larger scale irregularities and structures through transport effects. Quasi-linear transport theory yields diffusion coefficients which are functions only of the macroscopic parameters and may be used in fluid computer simulations. However, this theory should be generalized to include collisions and other sources of free energy, especially electric fields perpendicular to the magnetic field. Two dimensional forms of the diffusion coefficient should be incorporated into a realistic ionospheric structure model code.

In the high-latitude E region, the strong electrojet currents drive plasma instabilities, which lead to both small- and large-scale structure. The linear theories of these instabilities can explain several features of the irregularities. Results from nonlinear studies, of these instabilities (e.g., saturated amplitudes) are conflicting in several respects. Further analytical and numerical studies of these instabilities, especially the Farley-Buneman instability, are needed.

## EXPERIMENTAL EMPHASIS

The future direction of high-latitude experimental programs should benefit from lessons learned in the study of equatorial irregularities where a synergistic scientific approach (between experiment and theory) led to a successful understanding of major low-latitude geoplasma mechanisms. However, the high-latitude irregularities problem is substantially more complex and measurement requirements go beyond those applied at the equator. Thermal plasma temperatures, ion composition and charged-particle drift velocities must be measured to establish  $M_i$ ,  $T_e/T_i$ , relative drift velocities, and the cold plasma contribution to current systems. Full energy and pitch angle distribution should be determined, including the suprathermal regime ( $kT_e \approx 50$  eV). Electric and magnetic field measurements should have vector capabilities and density fluctuation measurements should include determination of phase velocities for waves in the drift, ionacoustic, and ion-cyclotron regimes. In all cases, temporal and spatial resolution should be optimized within available subsystem support capability and the simultaneity in acquisition of all known (and suspected) production, dissipation, and transport processes is a priority concern over the entire spectrum of scale sizes (hundred of kms to fractions of a meter). It is only through such complete diagnosis that any important, unidentified physical process can be discovered.

A critical scale size presently identified by theory is the ion gyroradius ( $\approx 3m$ ). Therefore, experiment and theory should provide a complete description of the wavelength dependence of each parameter which can influence the seeding, driving, dissipation, and transport processes down at least to this scale.

Continuing to emphasize complementarity in measurements, we note that the growth rate of the ExB gradient drift instability depends on the velocity difference between the ion and neutral gases. Thus, ion drift measurements must be complemented by simultaneous optical Fabry-Perot interferometer observations to assess the importance of this structuring process. The current-convective instability is driven by the current carried by thermal particles only. Thus, magnetometer measurements of total integrated current must be supplemented by instrumentation capable of determining that portion of the total current carried by the thermal plasma.

Another outstanding theoretical concern is the field-line coupling to other, perhaps distant, plasma processes and populations, e.g., E-region shorting and magnetospherically imposed structure. Thus, in situ measurements must span a large range of density, energy and pitch angle regimes and must be augmented by simultaneous diagnostics, such as incoherent scatter, which can measure vertical as well as horizontal plasma structure. These diagnostics should be applied to active experiments, e.g., conductivity, modification, to test specific hypotheses in controlled ways.

Finally, any viable theory must ultimately explain not only the amplitude of ionospheric structure, but its complex three-dimensional geometry. At present the only diagnostics that have some capability for specifying this geometry are incoherent scatter and spaced-receiver scintillation observations. However, even when combined, these instruments only span wavelengths from hundreds of kms to hundreds of meters. Thus, ultimately one must consider multiple in situ diagnostic systems (i.e., rocket satellite, and/or satellite-satellite multipoint measurements) to unfold the multitude of space- and time-dependent processes active in the high-latitude ionosphere.





# CHAPTER 1

## RECONNECTION OF MAGNETIC FIELDS

### WORKING GROUP MEMBERS

B. U. Ö. Sonnerup, Chairman  
*Dartmouth College*

P. J. Baum  
*University of California, Riverside*

J. Birn  
*Los Alamos National Laboratory*

S. W. H. Cowley  
*Imperial College, U.K.*

T. G. Forbes  
*University of St. Andrews, U.K.*

A. B. Hassam  
*University of Maryland*

S. W. Kahler  
*Emmanuel College*

W. H. Matthaeus  
*College of William and Mary*

W. Park  
*Princeton Plasma Physics Laboratory*

G. Paschmann  
*Max Planck Institute for Extraterrestrial Physics, F.R.G.*

E. R. Priest  
*University of St. Andrews, U.K.*

C. T. Russell  
*University of California, Los Angeles*

D. S. Spicer  
*Naval Research Laboratory*

R. Stenzel  
*University of California, Los Angeles*

## CHAPTER 1

### RECONNECTION OF MAGNETIC FIELDS

I.	Introduction	1-3
II.	Reconnection at the Magnetopause	1-7
III.	Reconnection in the Geomagnetic Tail	1-11
IV.	Particle Acceleration in Current Sheets	1-17
V.	The Diffusion Region	1-20
VI.	Laboratory Simulation of Magnetospheric Reconnection	1-26
VII.	Reconnection and the Evolution of Solar Fields	1-31
VIII.	Reconnection in Solar Flares	1-35
IX.	Acceleration of Solar-Flare Electrons by Reconnection	1-39
X.	MHD Simulation and Small Scale Structure	1-40
XI.	Laboratory Observations of Neutral Sheets	1-45
XII.	Reconnection in Tokamaks	1-52
XIII.	Conclusion	1-60
XIV.	References	1-62

## I. INTRODUCTION

(B. Sonnerup)

In cosmic plasmas, large scale lengths,  $L$ , large velocities,  $V$ , and small electrical resistivities,  $\eta$ , combine to form large values of the magnetic Reynolds number  $R_m = \mu_0 V L / \eta$ , a circumstance in which the plasma and magnetic field are tightly coupled or "frozen" together. If the ratio of plasma to magnetic energy density is large, nonuniform motions in such plasmas often stretch magnetic loops or push differently magnetized regions together into configurations where the magnetic field exhibits large shear, i.e., it changes direction and magnitude rapidly across a narrow electric current sheet. If the ratio is small, the magnetic field organizes the plasma motion instead and the currents have a tendency to flow along magnetic field lines, a situation that also leads to sheared magnetic fields. All such field configurations contain free magnetic energy. Magnetic field reconnection provides a means of converting some or all of this energy to plasma kinetic and thermal energy and thus of changing the field configuration towards one of thermal equilibrium.

The definition adopted by the working group is that magnetic field reconnection occurs in a plasma whenever an electric field  $E_{\parallel}$  is present along a magnetic separator, i.e., along a line of intersection of two separatrix surfaces which divide space into different magnetic cells as illustrated in Figure 1-1. The separator is also commonly referred to as the reconnection line, the X line or the null line. The electric field causes the plasma to drift across the separatrix surfaces from cells 1 and 2 into cells 3 and 4 carrying with it frozen magnetic field lines at a rate proportional to  $E_{\parallel}$ . In the process, the connectivity of these lines changes as illustrated in the figure and magnetic flux is transferred from the former two cells to the latter. This description of the process justifies the term "reconnection" but it should be emphasized that the definition given above does not depend on the use of the frozen-field condition, and that, indeed, this condition is violated at the separator.

If reconnection occurs in a nonsteady fashion, one expects the angle  $\theta$  between the separatrix surfaces in Figure 1-1 to increase as cells 3 and 4 fill up at the expense of cells 1 and 2, the maximum value being  $\theta = \pi/2$  for a curl-free field. But if the boundary conditions permit continual removal of plasma from the former two cells and continual supply to the latter two cells then a steady-state version of the process can be established. We shall refer to the former situation as "impulsive" or "spontaneous" reconnection and to the latter as "steady" or "driven" reconnection.

The earliest quantitative model of driven reconnection, developed by Sweet (1958) and Parker (1963), is shown in Figure 1-2a. Magnetic flux is carried towards a field-reversing current sheet of length  $L$  with a small speed  $v_1$  that is equal and opposite to the resistive diffusion speed. The plasma leaves the system by flowing along the current sheet away from the X line with a large speed  $v_2$  comparable to the Alfvén speed  $v_{A1}$ , the latter based on the plasma density and magnetic field in the uniform upstream field regions. The converted electromagnetic energy is carried away as kinetic energy of the plasma in this outflow. The reconnection rate is proportional to  $E_{\parallel}$  but for the steady case it is commonly expressed as the ratio  $M_A = v_1 / v_{A1}$ . For the Sweet-Parker model we have  $M_A \sim R_m^{-1/2}$ . This rate is extremely small in cosmic applications as a result of large  $R_m$  values.

A much larger reconnection rate was achieved in the model developed by Petschek (1964) and shown in Figure 1-2b and in the models by Yeh and Axford (1970) and Sonnerup (1970). In all of these, the so-called diffusion region, i.e., the narrow channel around the separator in which the current density is high and in which resistive and other effects help break the frozen-field

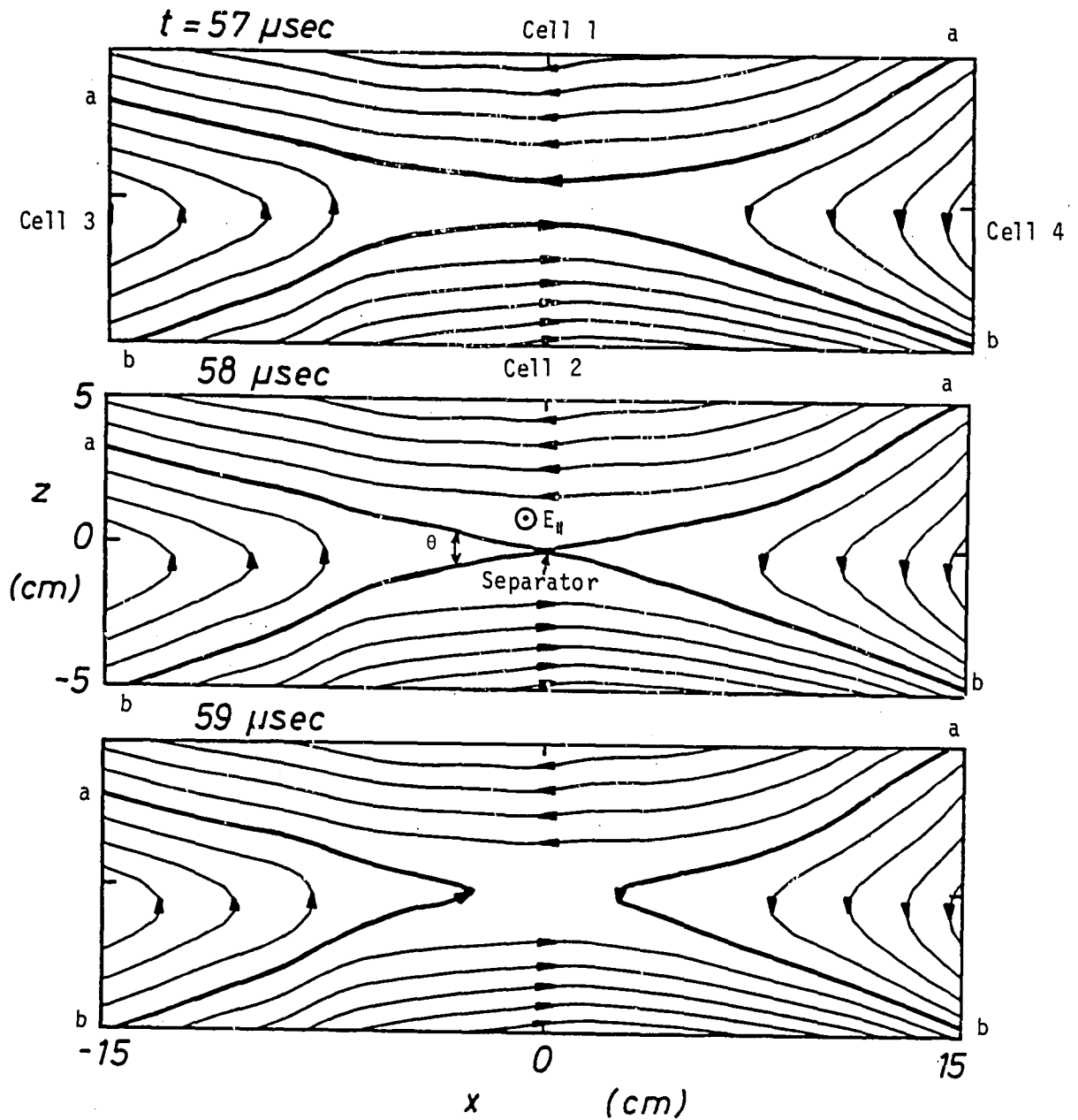


Figure 1-1. Evolution of magnetic field configuration in the laboratory reconnection experiment discussed in Section XI. The reconnection process is illustrated by the topology change of the pair of field lines  $a-a$  and  $b-b$  (heavy lines). During reconnection, plasma and magnetic flux are transferred from cells 1 and 2 to cells 3 and 4. The separator is located at the center of the figure and, along with the reconnection electric field  $E_{||}$ , is perpendicular to the plane of the paper.

condition, has been reduced in size from length  $L$  to length  $\ell \ll L$ . The Sweet-Parker model applies to this region but the magnetic Reynolds number based on  $\ell$  is now sufficiently small so that  $M_A$ , evaluated in the inflow adjacent to the diffusion region, is of order unity. In Petschek's model, the value of  $M_A$  far upstream is considerably smaller because the magnetic field is larger there. With the boundary conditions used by Petschek, this  $M_A$  has a logarithmic dependence on the magnetic Reynolds number,  $R_m$ , based on  $L$  and typically has values in the range 0.05-0.1. The actual length,  $\ell$ , and width,  $\delta$ , of the diffusion region, and the current density present in it, depend on the effective resistivity near the separator. The physical processes responsible for this resistivity in a collision-free plasma remain poorly understood.

Outside of the diffusion region, resistive effects are unimportant and the magnetic field and plasma are frozen together. In Petschek's model, the deflection and acceleration of the inflow velocity  $v_1$  to the outflow velocity  $v_2 \approx v_{A1}$  is achieved by the Maxwell stresses in pairs of MHD slow shocks. An asymmetric version of this model was also developed for magnetopause reconnection (Figure 1-2c) where the flow deflection and magnetic field reversal is accomplished in a rotational discontinuity (Lewy et al., 1964). In collision-free reconnection models (Hill, 1975) the plasma energization is also assumed to occur in a single current sheet.

The early stage of spontaneous reconnection is associated with the tearing instability (Furth et al., 1963). As illustrated in Figure 1-3, the tearing mode leads to the formation of an array of X-type and O-type neutral lines and associated magnetic islands in regions of sufficiently large field shear. Depending on the circumstances, the nonlinear evolution of this mode may either comprise an impulsive reconnection event or it may lead to steady reconnection. Another possibility is that the instability saturates at low amplitudes thus providing tearing turbulence but no large-scale reconnection in the current sheet.

In the reconnection configurations shown in the Figure 1-2, the magnetic field is confined to the plane of the paper and has opposite directions on the two sides of the current layer. But the process can operate in the presence of a magnetic field component perpendicular to the plane of the figures. Indeed, in a tokamak this is the dominant component. Its influence on the reconnection rate in a collision-free plasma is not well understood.

Magnetic field reconnection transcends the traditional disciplines of laboratory, space, and cosmic plasma physics. The process is known to occur in fusion devices such as tokamaks where it causes major disruption of the plasma confinement. It is thought to occur in solar flares and to be important in other less energetic events in the solar chromosphere and corona as well. And there is mounting evidence that reconnection takes place at planetary magnetopauses and in planetary, as well as cometary, magnetic tails. In all likelihood, reconnection also plays an important role in other astrophysical objects such as accretion discs as well as in a variety of current sheets occurring in interplanetary, interstellar, and intergalactic space. Detailed study of reconnection processes in the laboratory, in computer simulations, and in the Earth's magnetosphere is therefore an important task which will ultimately have a strong and positive impact on plasma astrophysics.

In the following sections, the reconnection process itself and some of the applications mentioned above will be discussed. The report is not a review article or even a complete statement of our current understanding of reconnection. But the individual topics presented provide a sense of the flavor and depth of some of the ongoing research as well as a summary of some of the outstanding unsolved problems.

In brief summary, the chapter is organized as follows. Recent spacecraft observations of steady and nonsteady reconnection at the magnetopause are reviewed in Section II. Computer simulations of three-dimensional reconnection in the geomagnetic tail are discussed in Section III.

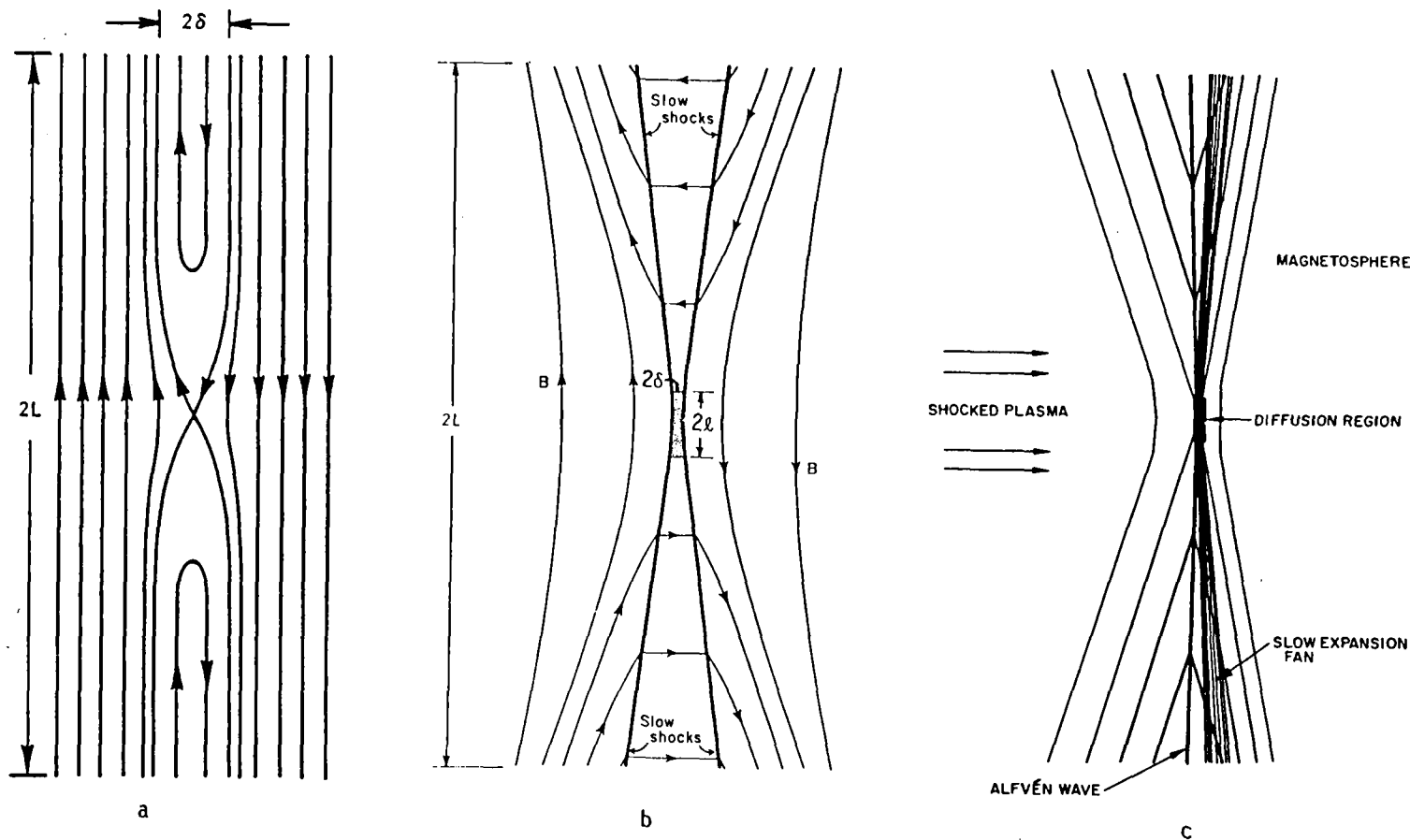


Figure 1-2. Two-dimensional steady MHD reconnection models (after Petschek, 1964; 1966).

- (a) Sweet-Parker model. Slow plasma inflow takes place from the left and right with rapid outflow ( $v_2 \approx v_{A1}$ ) at top and bottom in a layer of width  $2\delta$ .
- (b) Petschek model. Slow plasma inflow from both sides. The plasma is accelerated into the wedge-shaped outflow regions by slow shocks and leaves the system at top and bottom with  $v_2 \approx v_{A1}$ . Diffusion region at center has height  $2\ell$  and width  $2\delta$  where  $\delta < \ell \ll L$ .
- (c) Levy et al. (1964) magnetopause reconnection model. Plasma inflow from the left only, with vacuum conditions in the magnetosphere on the right. Plasma acceleration into the outflow,  $v_2 \approx 2v_{A1}$ , at top and bottom is accomplished in a large-amplitude Alfvén wave or rotational discontinuity.

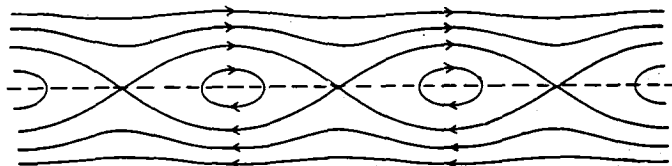


Figure 1-3. Tearing mode magnetic field configuration.

Section IV and V present theoretical aspects of the energization of particles in current sheets and of the microprocesses in the diffusion region, respectively. Section VI describes terrella experiments in which magnetospheric reconnection, both at the magnetopause and in the tail, is simulated. The report then turns to a discussion of the possible role of reconnection in the evolution of solar magnetic fields, in Section VII, and in solar flares, in Sections VIII and IX. Two dimensional MHD computer simulation of turbulent reconnection with possible application to coronal heating is examined in Section X. Section XI contains a review of results from an ongoing laboratory reconnection experiment which has possible solar as well as magnetospheric application and, in Section XII, recent results concerning reconnection in tokamaks are presented. The final section of the report highlights scientific results and problems concerning reconnection that have emerged from the study.

## II. RECONNECTION AT THE MAGNETOPAUSE

(G. Paschmann and C. T. Russell)

Over the years, a great deal of observational information has been accumulated, indicating that reconnection as described by Dungey (1961) (see Figure 1-4) is an important process both at the magnetopause and in the tail of the Earth's magnetosphere. In this framework, interplanetary magnetic field lines, carried towards the magnetopause by the shocked solar wind, reconnect with terrestrial field lines in the diffusion region surrounding the separator. These field lines are then pulled over the poles into the magnetotail where they eventually reconnect at a second reconnection site and become closed again. On the dayside, the shortening of the reconnected field lines implies conversion of electromagnetic energy into plasma energy. Tailward of the cusp, field lines are stretched out and electromagnetic energy is extracted from the solar wind flow. Studies of reconnection naturally focus on the dayside magnetopause where the process is initiated, even though the transfer of energy from the solar-wind plasma to the magnetosphere occurs mainly in the tail magnetopause.

The earliest evidence of magnetopause reconnection was provided by the dependence of geomagnetic activity on the north-south component of the interplanetary magnetic field. More direct evidence is the sensitivity of the location of magnetospheric boundaries on the north-south component of the IMF and the appearance of flux in the geomagnetic tail. In rare instances direct in situ evidence in the form of a significant component  $B_n$  of the magnetic field normal to the magnetopause has been deduced from the measurements (Sonnerup and Ledley, 1979). But, for lack of adequate instrumentation, the plasma and electric field signatures of magnetopause reconnection have remained undetected until recently.

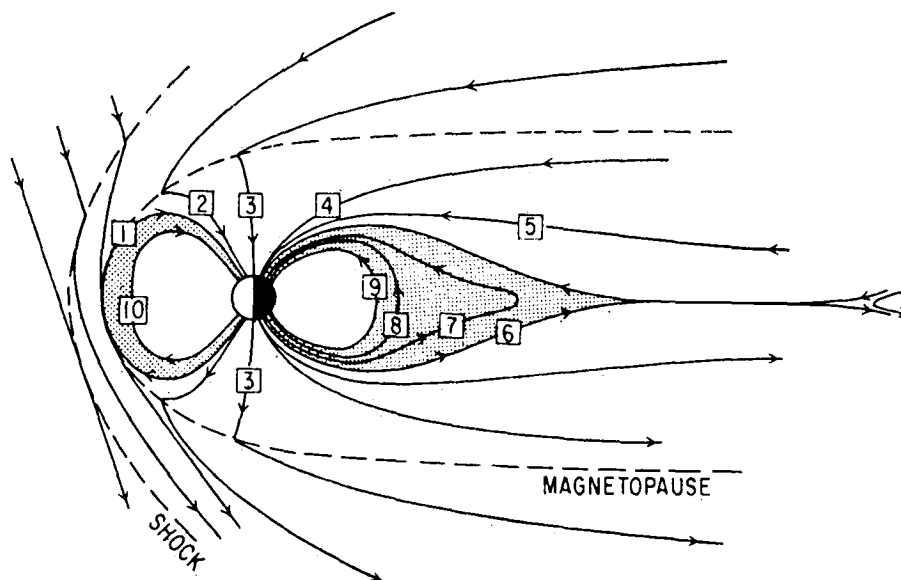


Figure 1-4. Meridional view of the magnetosphere undergoing steady reconnection at the magnetopause and in the geomagnetic tail. The numbers indicate the successive positions of magnetic field lines. (After Axford, 1969.)

Since 1977, the ISEE mission, with its much improved measurements (e.g., three-dimensional plasma flows; electric fields), high time resolution, and two-spacecraft aspect, has provided in situ evidence for reconnection both in a quasi-steady form and in the form of patches caused by the intermittent localized occurrence of the process. Outlined below are some of the experimental findings and outstanding problems relating to both these forms of magnetopause reconnection.

### QUASI-STEADY MAGNETOPAUSE RECONNECTION

The theoretical model developed by Levy et al. (1964) and shown in Figure 1-2c describes an asymmetric version of reconnection which is expected to occur when the cold, dense solar wind plasma approaches the magnetosphere with its hot, dilute plasma. In this model the magnetopause appears as a rotational discontinuity in which the magnetopause experiences a tangential velocity increase  $\Delta v_t$  which is equal to the Alfvén speed based upon the change in tangential magnetic field,  $B_t$ , i.e.,

$$\Delta v_t = \frac{\Delta B_t}{\sqrt{4\pi\rho}}.$$



This acceleration is the result of the  $\underline{l} \times \underline{B}_n$  force within the magnetopause where  $\underline{l}$  is the magnetopause current. It is directed away from the separator. For typical densities and magnetic field strengths, and for field rotation angles of  $\sim 180^\circ$ ,  $\Delta v_\perp$  is of the order of 300 km/sec. Note that the velocity change can also be derived on the basis of single-particle considerations (see Section IV).

So far, twelve instances of high-speed flows at the magnetopause have been reported which obey the above equation within the uncertainties of the measurements (Paschmann et al., 1979; Sonnerup et al., 1981; Gosling et al., 1982). The duration of these events was sometimes large enough for the process to be considered nearly stationary. For some of these events, additional support for the reconnection interpretation has been obtained from observed energetic particle anisotropies from which it could be inferred that magnetospheric field lines indeed crossed the magnetopause (e.g., Scholer et al., 1981).

The question of energy balance in some of the high-speed flow events has been addressed only recently: can the electromagnetic energy converted during reconnection be accounted for in terms of increased kinetic energy and temperature of the plasma flowing across the magnetopause, and in terms of heat flowing away from it? To check the energy balance, one must deal with up to the third moment of the measured distribution functions and associated large experimental uncertainties. Nevertheless, it appears that energy balance is maintained and that both kinetic-energy and temperature increases contribute to the balance.

Electric field measurements on ISEE have provided independent evidence for magnetopause reconnection. Although the tangential electric fields are expected to be typically around 1 mV/m, which is close to the sensitivity limit of the electric field probes, there have been a number of positive identifications. (Mozer et al., 1979; Formisano et al., 1982; Aggson et al., 1983). Most recently (Aggson, private communication), electric field measurements have directly confirmed the interpretation of several of the high-speed flow cases analyzed by Sonnerup et al. (1981).

It should be noted that in all the cases described above the evidence for reconnection was derived from the observation of one or more of the direct consequences of the process. The direct measurement of an electric field along the separator, ideally required as a proof (see Section I), is difficult, if not impossible, to make in the case of the magnetopause. Nevertheless, in total, one is justified in stating that there is very good evidence from *in situ* measurements that quasi-steady reconnection occurs at the magnetopause. Yet a number of problems and questions remain.

Reconnection Rate and Scale Size. Measurements of high-speed flows are not in themselves sufficient to define the reconnection rate (see Section IV). Electric-field data are also difficult to translate into a reliable reconnection rate, in part because the largest  $\underline{E}$  component is usually perpendicular to the magnetopause so that an accurate magnetopause normal vector is needed. Furthermore, the size of the region over which reconnection occurs at any given time is not known. For these reasons, it is presently difficult to assess with confidence the contribution of quasi-steady reconnection at the magnetopause to the overall magnetospheric dynamics.

Onset Criteria. There is some evidence that the magnetopause (at least locally) can be a tangential discontinuity even if the IMF is oriented such that the magnetic field reverses direction at the magnetopause. This may be a manifestation of the small-scale size of the process. It might, however, also mean that the appropriate interplanetary field direction is only a necessary but not a sufficient condition for the onset of reconnection.

Properties of the Rotational Discontinuity. The observed density variations within the magnetopause during some of the high-speed flow cases require the plasma to be anisotropic because  $\rho(1-\alpha) = \text{const.}$  across a rotational discontinuity, where  $\alpha \equiv 4\pi(p_{\parallel} - p_{\perp})/B^2$ . The ISEE plasma data

do not allow the validity of this relation to be checked adequately. More importantly, the rotational discontinuity is observed to be dissipative, a property not predicted (but also not excluded) by existing theory.

Diffusion Region. A major deficiency of the ISEE results obtained to date is the fact that no encounter with the diffusion region has yet been convincingly demonstrated. This means that the microphysics of the reconnection process has not been part of the investigation. The lack of reported encounters with the diffusion region might be attributed to its small expected size, but it is also possible that a lack of precise knowledge of what signatures to expect (see Section V) might have prevented its identification.

## FLUX TRANSFER EVENTS

The reconnection signatures described above are observed more often than might be inferred from the fact that only twelve cases have been analyzed so far. However, a more common response of the magnetopause to southward interplanetary fields appears to be time-varying reconnection of single flux tubes with  $\sim 1 R_e$  diameter. Figure 1-5 shows a schematic of such an event. Reconnection takes place over a limited region and for a limited time and then connected tubes are convected to the north and south. The most prominent signature in the magnetic field observed near the magnetopause is a large wiggle in the field component normal to the magnetopause. For northward motion of the flux tube, the unconnected field surrounding the tube is first deflected outward and then inward. For southward motions, the sequence is reversed. Early in the ISEE mission, only the former signature was observed. Since these events had the characteristics of magnetic connection between the magnetosheath and the magnetosphere and since the structures appeared to be convected with the magnetosheath plasma, they were termed flux transfer events, or FTEs (Russell and Elphic, 1978). Subsequently, also the reverse signature was discovered (Rijnbeek et al., 1982; Berchem and Russell, 1983). The distribution of the two signatures is basically consistent with a separator which passes through the subsolar point on the magnetopause. FTEs seem to occur whenever the interplanetary field is directed southward ( $B_z < 0$ ); they are almost never observed for  $B_z > 0$ . Since FTEs carry a significant field-aligned current, they might leave noticeable footprints in the ionosphere, which have not yet been looked for.

The evidence for magnetic connection in FTEs comes primarily from the observation of magnetospheric energetic particles streaming outward along the flux tubes (e.g., Daly et al., 1981; Scholer et al., 1982). Another indication is that magnetosheath plasma is seen whenever FTEs are detected inside the magnetosphere. Many FTEs show intense electron heat flux layers at the edges of the flux tubes (Scudder et al., 1983). This observation indicates that in this class of events, energy is still being dissipated as if reconnection was still going on. The possibility that these heat fluxes originate in the diffusion region should be examined.

Support for the interpretation of FTEs as illustrated in Figure 1-5 also comes from the consideration of the normal stress balance. Paschmann et al. (1982) have examined the overpressure in connected flux tubes and find that it is equal to the Maxwell stress imposed by the draping of the magnetic field around the tube. The fact that flux transfer events consist of twisted flux helps maintain their 3-D character and prevents them from being flattened against the magnetopause.

In summary, flux transfer events seem to play an important role in the dynamics and maintenance of the magnetopause boundary layer and might actually be the dominant mode of magnetopause reconnection (Cowley, 1982). The lack of theoretical guidance for this type of three-dimensional time-dependent reconnection, however, makes further progress in the analysis of the events difficult.

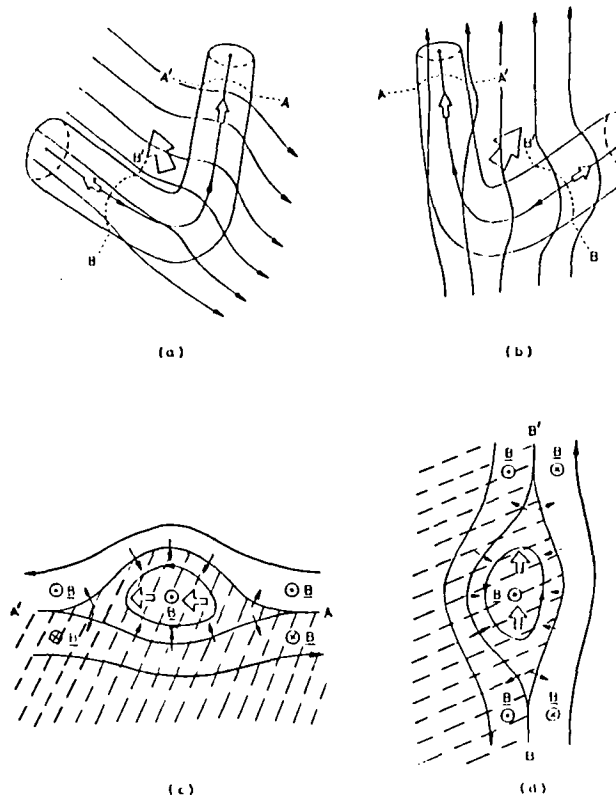


Figure 1-5. Northern half of a flux transfer (a) seen from the magnetosheath, (b) seen from the magnetosphere, (c) in cross section along A-A', and (d) in cross section along B-B'. Flux transfer events result from the temporally and spatially limited reconnection of magnetospheric and magnetosheath magnetic field. Twisting of field lines in the flux transfer event preserves their cylindrical shape. Large arrows denote flux tube motion. Flux tube diameters are approximately  $1 R_E$ . (After Cowley, 1982.)

### III. RECONNECTION IN THE GEOMAGNETIC TAIL

(J. Birn)

Reconnection at the magnetopause must on the average be balanced by an equal amount of reconnection in the geomagnetic tail, otherwise magnetic flux would continually be added to the two tail lobes without a corresponding flux removal at the neutral sheet separating the two lobes. However, the process does not necessarily take place simultaneously or at the same rate at the two sites. One possible model of tail reconnection which has emerged from an extensive study of plasma and magnetic field observations (Hones, 1980) is shown in Figure 8-7 of Chapter 8. Initially, a small amount of reconnection may be present at a reconnection line in the distant tail. As the interplanetary magnetic field turns southward, increased magnetopause reconnection leads to flux addition to the two tail lobes and an associated storage of free magnetic energy. The observed changes in the tail configuration include gradual thinning of the plasma sheet and field line stretching which ultimately lead to the onset of the tearing instability with a new neutral line forming at a point in the neutral sheet located in the vicinity of the midnight

meridian and much closer to Earth than the original reconnection line. Two models for the onset of the tearing instability are described in Section V. As this instability develops nonlinearly and the longitude extent of the reconnection line increases, a magnetic bubble or a plasmoid is formed and ejected tailwards. Extensive observations that support this reconnection model are discussed in detail in Chapter 8. As in the case of the magnetopause, the most convincing evidence for tail reconnection is the observation of plasma jetting with speeds of the order of the Alfvén speed. In this case, the observed jets may be directed either earthwards or tailwards depending on the spacecraft location and on the phase of the event.

In the last few years, magnetotail reconnection has been investigated extensively by use of numerical models. Early calculations have shown that tearing instability in an initially plane sheet with periodicity along the sheet proceeds slowly and does not accelerate the plasma to speeds comparable to the Alfvén speed. Several modifications have been studied to overcome this problem:

(1) Possible coalescence of tearing-mode magnetic islands.

(2) Modes driven by plasma streaming and associated non-uniform electric fields (Sato et al., 1983).

(3) Tearing modes in asymmetric tail-like configurations (Birn and Hones, 1981) where the mode evolves into impulsive reconnection without being externally driven.

Although the studies (2) and (3) are different, they have produced similar features, consistent with the basic reconnection picture and/or with substorm associated effects observed in the tail:

(a) Plasma is expelled from an X-type reconnection region at speeds of the order of the Alfvén speed; a typical flow pattern is generally found close to the X-line such that earthward flow is connected with positive  $B_z$ , and tailward flow with negative  $B_z$ , using the GSM coordinate system with the Z direction perpendicular to the plasma sheet; there are, however, modifications to this picture as discussed below.

(b) At some distance from the center plane this flow becomes field aligned.

(c) Two current layers form inside the magnetic separatrix that are similar to the slow shocks of Petschek's (1964) theory (see Figure 1-2b).

(d) Plasma sheet thinning; in model (2) this results from the driving force, whereas in model (3) it is a consequence of the spontaneous dynamic process.

(e) The reconnection region remains limited in the GSM Y direction across the tail, its extent depending on the initial configuration in model (3) and on the non-uniform boundary condition with the applied electric field a maximum in the center in model (2). The cross-tail current tends to avoid flowing through the reconnection region.

(f) Field-aligned currents are found both in the driven reconnection of model (2) and in the spontaneous reconnection of model (3). Although the general pattern is similar in both approaches, the directions of the field-aligned currents are opposite, corresponding to region 1 (earthward currents on the dawn side, tailward currents on the dusk side) in the former case and to region 2 in the latter case.

In addition, a number of detailed 3-D and time-dependent features have been predicted in the asymmetric tail model (3):

- (g) The formation and tailward expulsion of a plasmoid (see Figure 1-6).
- (h) A complex reconnection region with multiple neutral lines, connected in different ways; however only one X-line shows the typical reconnection flow pattern (see Figure 1-7).
- (i) The magnetic neutral points,  $B=0$ , and the flow stagnation points,  $v=0$ , do not coincide: in a nonsymmetrical nonsteady situation there is indeed no reason why they should. The region of negative  $B_z$  is generally more limited in space and time than the region of strong tailward flow.
- (j) The magnetic field pattern of the reconnection region is closer to a Y-type than to an X-type magnetic null with strong positive  $B_z$  earthward of the main neutral line and weak, on the average, negative  $B_z$  tailward of that line.

The numerical studies performed so far have produced a number of important questions and controversies. Among these are:

Driven Versus Spontaneous Reconnection: Observations suggest that the occurrence of substorms is closely related to the solar-wind electric field component  $E_y$ , whereas the dynamics of an individual substorm seems not to be controlled by the solar wind. This suggests that a driving mechanism is necessary to initiate reconnection but that the dynamics of this process is controlled internally. It is an open question, to what extent the growth of a spontaneous tearing mode is altered by the presence of a driving mechanism. Computer studies in a realistic geometry with an external electric field applied to the tail seem desirable.

Field Aligned Currents: The apparently contradictory results concerning the direction of field-aligned currents from models (2) and (3) could constitute different parts of a total picture, where region 1 currents are driven externally and region 2 currents result from the internal dynamics. More simulations seem desirable.

Particle Acceleration: If the reconnection process were entirely driven, the electric potential drop along the neutral line would be determined by the external electric field. This is insufficient for the acceleration of particles to energies of the order of, or greater than, 1 MeV during a single traversal of a particle across the reconnection region because the typical total potential difference  $\Delta\phi < 100$  keV. Although the impulsive reconnection of model (3) could lead to larger induced electric fields the present models show a maximum potential difference of similar magnitude, perhaps as a result of the unrealistic constant resistivity used in the simulation. A more realistic resistivity, increasing with increasing current density, might produce explosive reconnection with sufficiently high  $E_y$ . Another alternative is suggested by the fact that the region tailward from the main neutral line has the character of a thin neutral sheet in which the electric field magnitude and direction varies with space (as shown in Figure 1-8) and time, and by the fact that this region is limited in the Y direction. In such a configuration the possibility arises that particles may cross the neutral sheet several times in the Y direction being returned by the stronger positive  $B_z$  outside this region. Particles moving in the right phase could then be accelerated to energies equal to several times the potential drop. Calculations of particle orbits in these variable fields are desirable.

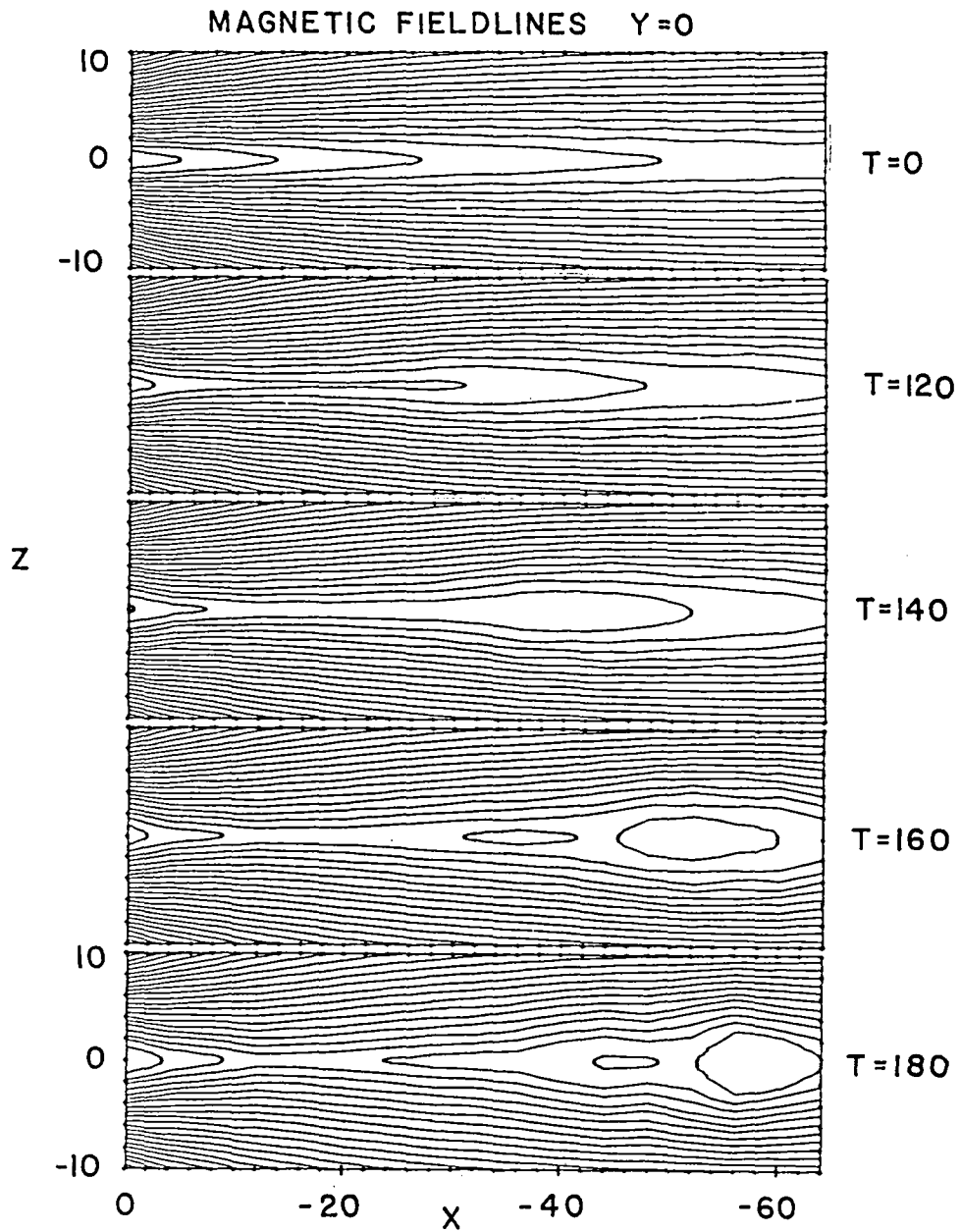


Figure 1-6. Evolution of magnetic fieldlines in a computer simulation of reconnection in the geomagnetic tail in the midnight meridian plane  $Y=0$ . The Earth is to the left outside the system. Length units are normalized by the initial plasma sheet half-width  $L_z$  at the left boundary and time units by a characteristic time  $T_A = L_z/V_A$ , where  $V_A$  is a typical Alfvén speed.

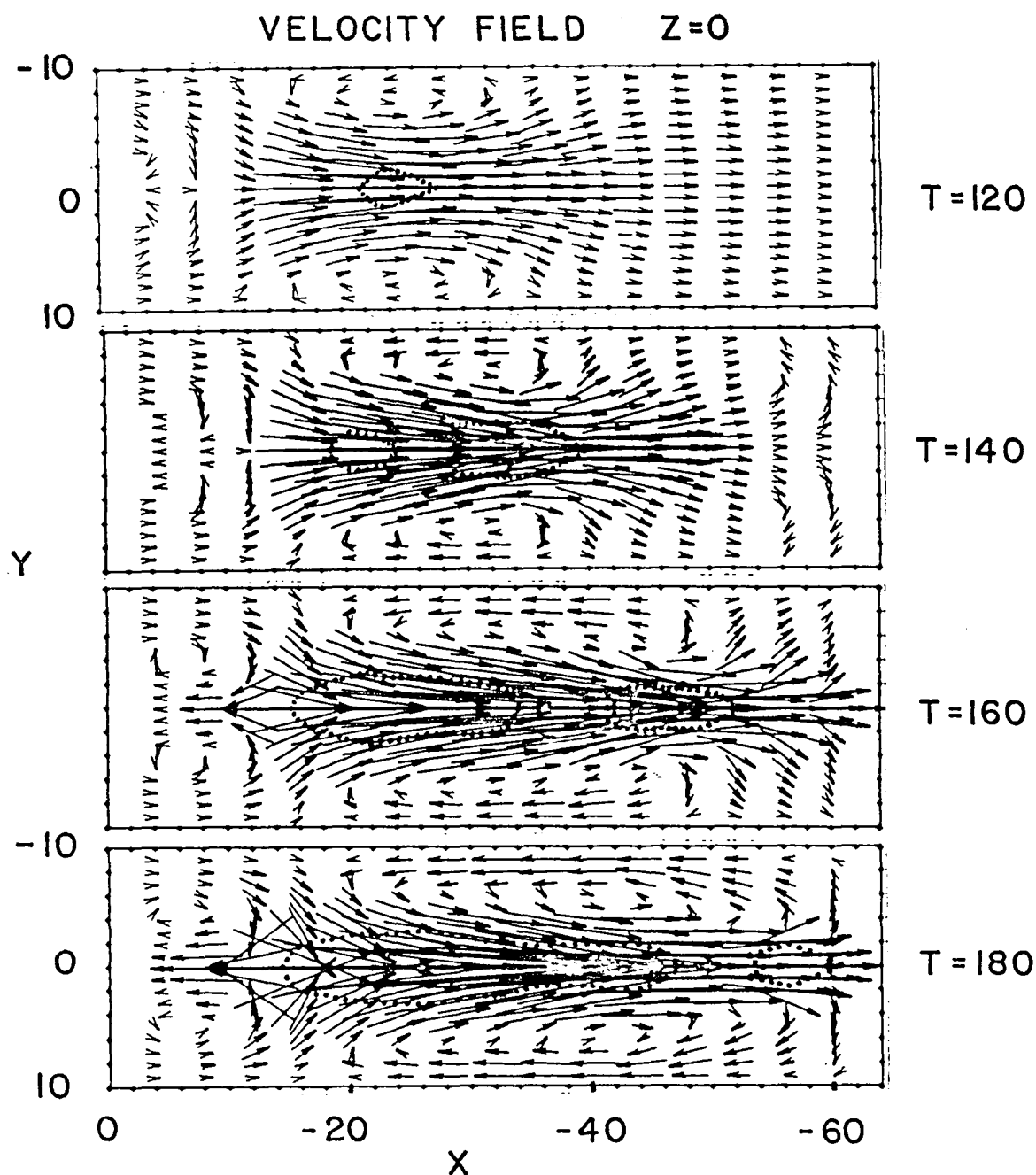


Figure 1-7. Evolution of the velocity field and the magnetic neutral line (dotted lines) in the equatorial plane  $Z=0$  for the same simulation as in Figure 1-6. The maximum length of velocity vectors corresponds to about  $1.2 V_A$ .

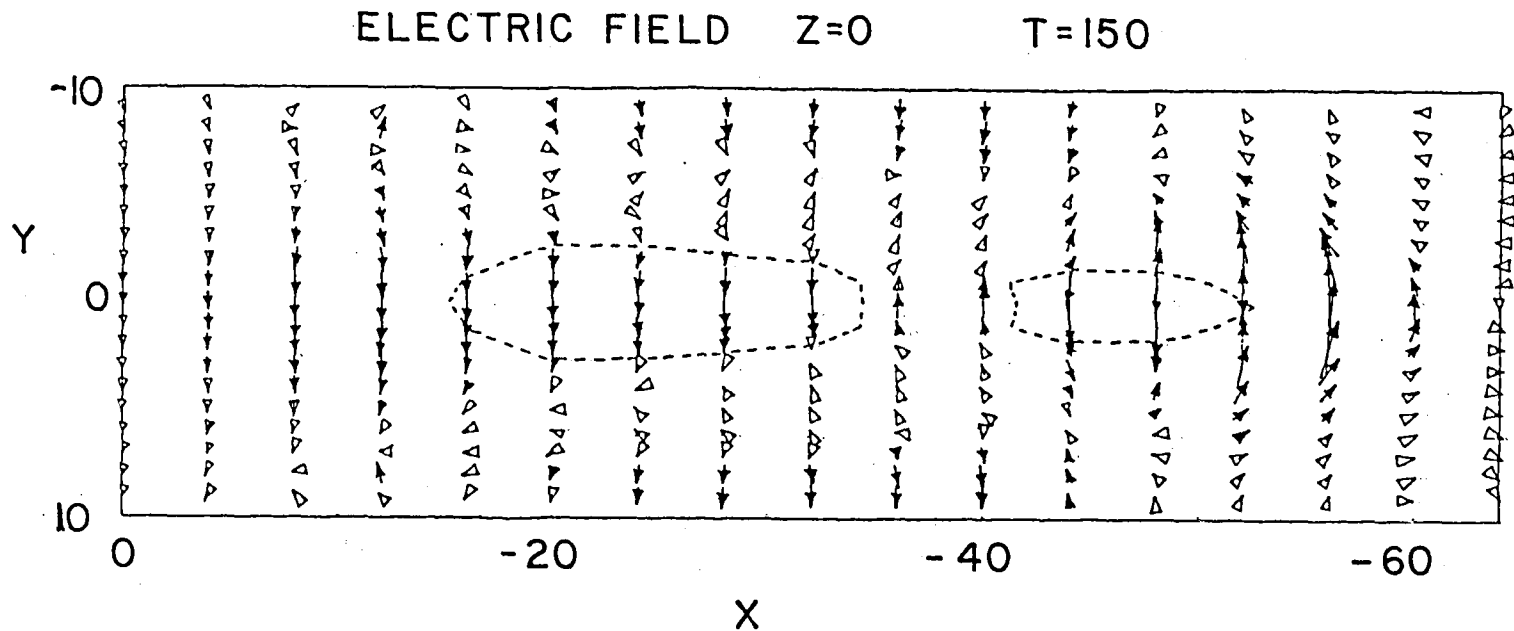


Figure 1-8. Electric field vectors and magnetic neutral line (dashed line) in the equatorial plane at  $t=150$  for the same simulation as in Figures 1-6 and 1-7. The maximum length of electric field vectors corresponds to about 1 mV/m assuming  $V_A=1000$  km/s and a lobe magnetic field strength of  $20\gamma$  at  $x=-20$ .



## IV. PARTICLE ACCELERATION IN CURRENT SHEETS

(S. W. H. Cowley)

An important aspect of the reconnection process is the acceleration of plasma particles by the electric field in the vicinity of the magnetic separator line and in the current sheets which lie downstream thereof. In general a component of the electric current density  $\underline{j}$  will flow in the direction of the reconnection-associated electric field  $\underline{E}$  and energy will hence be liberated to the plasma at the rate  $\underline{j} \cdot \underline{E}$  watts  $\text{m}^{-3}$ . In the steady state this energy is supplied by inflow of electromagnetic energy to the current sheets given by the Poynting flux  $\underline{S} = \underline{E} \times \underline{H}$  watts  $\text{m}^{-2}$ . The first quantitative estimates of the results of this energy input were made by Parker (1963), who showed within the framework of MHD that the plasma would be ejected from the current sheet region downstream of the separator with a speed comparable with the Alfvén speed of the inflow plasma  $V_A$  (see Figure 1-2a). MHD reconnection models have subsequently been extended and elaborated by numerous authors (e.g., Petschek, 1964; Sonnerup, 1970; Vasylunas, 1975) but ejection of the plasma at  $\sim V_A$  remains an invariant feature, since the result rests essentially on conservation of energy in the steady state between electromagnetic energy flowing into the current sheets and accelerated plasma flowing out. In particular, within limits the result does not depend upon the actual rate of reconnection, expressed for example in terms of the magnitude of the reconnection-associated electric field. This determines only the quantity of matter per unit time which is accelerated to  $\sim V_A$ .

Over the same time interval that fluid models of reconnection were being developed, complementary studies of the acceleration of particles applicable to collision-free plasmas were also undertaken by examining the motion of individual particles in model electromagnetic fields. In the vicinity of the magnetic separator, a neutral sheet magnetic field may be taken as a first approximation, i.e., the field is given by  $\underline{B} = (B_x(z), 0, 0)$  where  $B_x$  is an odd function of  $z$  and goes to zero at  $z=0$ . With the inclusion of a reconnection-associated electric field  $E_y$  in the direction of the sheet current, the motion of the particles is as shown in Figure 1-9a. Ions and electrons  $\underline{E} \times \underline{B}$  drift into the neutral sheet region from both sides and are then freely accelerated along the sheet by  $E_y$ , ions and electrons in opposite directions (Speiser, 1965; Sonnerup, 1971; Cowley, 1973; Stern, 1979). The energy that the particles gain in the system is not arbitrary, however, but is governed by the conditions of self-consistency. For a strict neutral sheet system of finite but arbitrary width Alfvén (1968) showed that the electric current carried by the accelerating particles satisfies Ampère's law if the voltage across the neutral sheet has the value

$$\Phi_A = \frac{B^2}{\mu_0 n e}$$

where  $B$  and  $n$  are field strength and ion/electron number density, respectively, in the inflow region outside the neutral sheet. The quantity  $\Phi_A$  is generally referred to as the Alfvén potential. The average ion accelerated essentially from rest through the potential  $\Phi_A/2$  exits the neutral sheet with the speed  $V_A$ , just as in the fluid models, but now in the direction of  $E_y$  and the current rather than at right angles to that direction. The average electron accelerated through the same potential exits at a speed  $(m_i/m_e)^{1/2}$  times higher at the opposite boundary to the ions. It is instructive that in Alfvén's system the current is limited solely by the finite residence time of ions and electrons in the neutral sheet before they reach the boundaries, corresponding to Speiser's (1970) "inertial conductivity." In principle, anomalous resistance is not required to limit the current or to allow reconnection to proceed, although clearly unstable distributions will be readily produced in these systems leading to microinstabilities which may be dynamically significant.

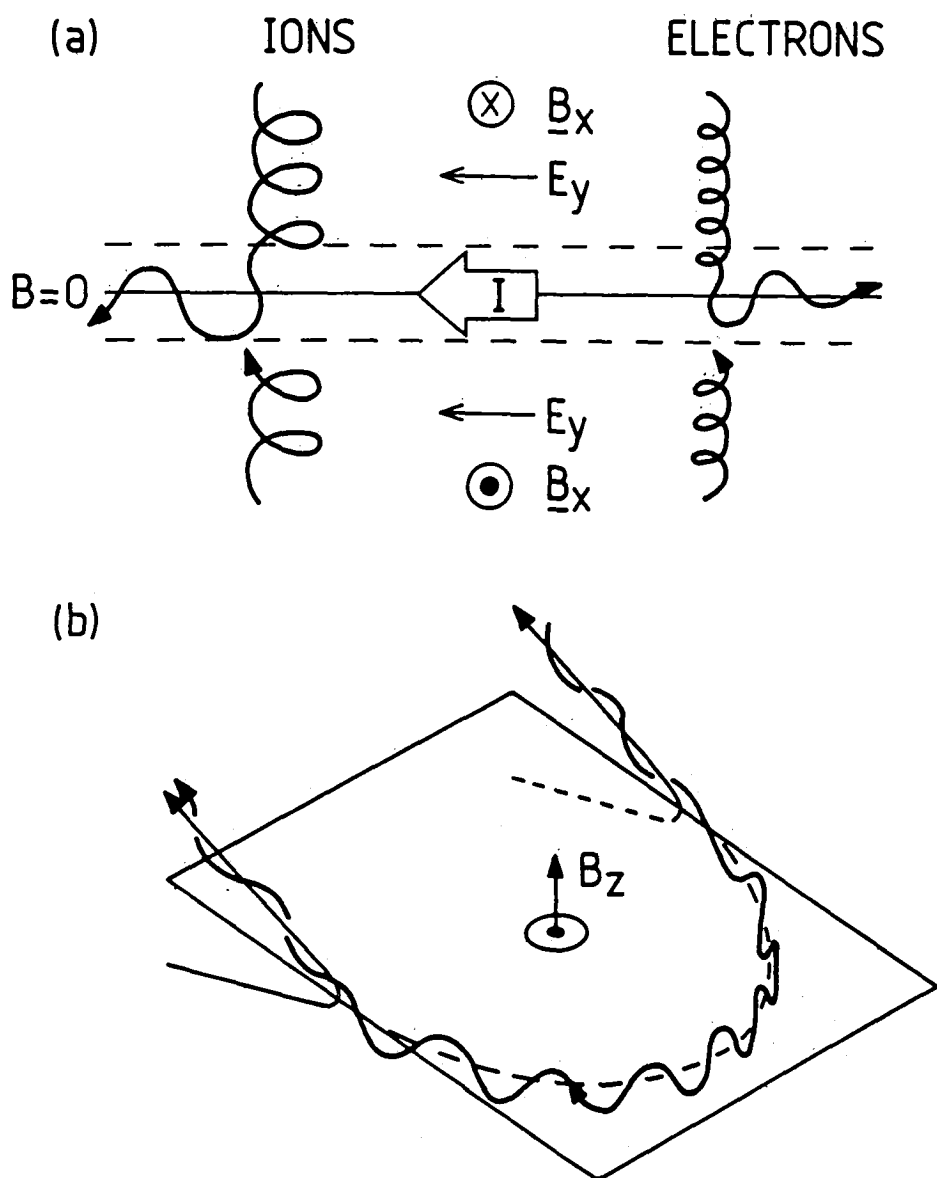


Figure 1-9. (a) Particle motion in a neutral sheet system with a uniform electric field in the direction of the current  $I$ . (After Speiser, 1965.)  
 (b) Perspective view of non-adiabatic "Speiser" ion motion in a current sheet field with finite uniform  $B_z$ , in the frame in which the electric field  $E_y$  has been transformed away. In the sheet the particle oscillates rapidly in the  $z$  direction due to the reversal of the  $B_x$  field, as in the neutral sheet, and also slowly turns a half circle of radius  $v/\Omega_z$  in the plane of the current sheet due to the Lorentz force associated with  $B_z$ . (After Dungey, 1953; Speiser, 1965.)

Away from the immediate vicinity of the separator a finite field  $B_z$  must be added to the neutral sheet field of Figure 1-9a, and this leads to an important modification of the particle motion. After drifting into the current sheet the particles begin to accelerate along it due to  $E_y$  as before, but now the Lorentz force associated with the  $B_z$  field causes the particles to turn in the X-direction. For an 'O' type magnetic field configuration the particles are turned toward the neutral line on both sides so that the particles remain confined to its vicinity as for the strict neutral sheet. However, for an X-type null the particles turn away from the neutral line on both sides and eventually leave the current sheet, flowing out along the reconnected field lines just within the separatrix (Dungey, 1953; Speiser, 1965, 1967). The cross-system displacement of these particles within the current sheet determined by  $B_z$  may only be a small fraction of the full width of the system. Since the current provided by the particles is then correspondingly reduced compared with the neutral sheet, the voltage across the current sheet required for self-consistency is not limited to  $\Phi_A$  but may be much larger. This effect corresponds to Speiser's (1970) "gyroconductivity" which may in fact be the dominant current-limiting process both in the immediate vicinity of the separator as well as in the down-stream region. The "inertial resistivity" of the strict neutral sheet which requires particles to travel the full width of the current sheet without being ejected by the  $B_z$  field will only be applicable in general if  $B_z$  is very weak for extended distances away from the separator.

To examine the behavior of particles and their acceleration downstream of the separator in the region where  $B_z$  is fairly uniform it is useful to adopt a "one-dimensional current sheet" approximation for the magnetic field  $\underline{B} = (B_x(z), 0, B_z)$ , with  $B_z$  a strict constant. The reconnection-associated electric field  $E_y$  may then be locally transformed away by moving to a frame having a speed  $V_T = E_y/B_z$  relative to the separator rest frame. This provides great simplicity since the particle speed in this frame is constant, allowing discussion of the acceleration process without the need for a detailed discussion of the particle motion in the current sheet. Nevertheless it is worth pointing out that with finite  $B_z$  the form of this motion depends strongly on the thickness of the equilibrium current sheet and will, e.g., be adiabatic for a sufficiently wide current layer (Francfort and Pellat, 1976; Cowley and Pellat, 1979). The more likely situation, however, is that the sheet thickness will have ion Larmor radius scales, such that the ion motion at least is non-adiabatic and has the form of "Speiser" orbits illustrated in Figure 1-9b in the  $E_y=0$  frame (Eastwood, 1972, 1974). The acceleration experienced by the particles in either case may be obtained using the condition for current sheet equilibrium together with a frame transformation argument (Cowley, 1980; Cowley and Shull, 1983). Assuming that the plasma flowing along the field lines into the current sheet in the  $E_y=0$  frame has similar properties on either side of the sheet so that there is no net plasma flow through the layer, the equilibrium condition is that the marginal firehose stability condition must just be met in the uniform field region outside the current layer (Rich et al., 1972). In the simplest case of cold, field-aligned streams of density  $n$  and speed  $v$  incident upon the layer from either side, this condition determines the stream speed in the  $E_y=0$  frame required for equilibrium as

$$v = \frac{B}{\sqrt{2\mu_0 n m_i}} = V_A$$

where  $V_A$  is again the Alfvén speed outside the current layer. The electron contribution to the stress balance (and hence the sheet current) is typically  $(m_e/m_i)$  times the ion contribution in this case and is hence negligible. If we now transform back to the neutral line frame the field-aligned inflow and outflow speeds of the ion streams will be given by (assuming  $B_x \gg B_z$  outside the current layer)

$$v_{in} = V_A - v_T \text{ and } v_{out} = V_A + v_T$$

so that in terms of  $v_{in}$ , assumed given, the required "field line speed"  $v_T$  and the outflow speed  $v_{out}$  are given by

$$v_T = V_A - v_{in} \text{ and } v_{out} = 2V_A - v_{in}$$

When  $v_{in}$  is small, therefore, the outflowing accelerated ions will form a nearly field-aligned stream with speed  $\sim 2V_A$ . In the geomagnetic tail this stream occupies the same region as the slowly moving inflow plasma so that the average velocity is  $V_A$  in agreement with the fluid model. At the magnetopause the inflow and outflow are usually on opposite sides of the current sheet so that the average outflow speed remains  $2V_A$ , again in agreement with the fluid model (compare Section II). Thus compatible results are obtained using the two approaches provided self-consistency conditions are applied to the single-particle picture.

Turning briefly to observations, it is to be emphasized that the Earth's magnetosphere provides unique opportunities for detailed in situ observations of reconnection-associated plasma acceleration. Jaeger and Speiser (1974) were the first to suggest that the accelerated ion streams observed in the geomagnetic tail during substorms (e.g., Hones, 1973, 1980; DeCoster and Frank, 1979; Coroniti et al., 1980) were associated with current sheet acceleration, while Cowley (1980) and Cowley and Southwood (1980) showed that the observed ion bulk speeds of  $\sim 500\text{--}1500 \text{ km s}^{-1}$  were compatible with the self-consistent 1-D current sheet model discussed above. They also were the first to discuss the spatial properties of the streams and their relationship with observed plasma populations, principally the plasma sheet. Similar, but much more detailed stress-balance considerations have been given to observations of reconnection at the dayside magnetopause by Sonnerup et al. (1981). At a distribution function level, Lyons and Speiser (1982) have shown that the peculiar velocity distributions of the ion jets observed in the tail can be successfully reproduced by current sheet acceleration. An example observed by ISEE2 on the outer boundary of the plasma sheet is shown in Figure 1-10, taken from Forbes et al. (1981).

Finally, it should be mentioned that although the major energy input from the electromagnetic field in the geomagnetic tail goes into 1-10 keV ions as described above, a high-energy tail population of both ions and electrons has also been observed extending to  $\sim \text{MeV}$  energies at disturbed times (e.g., Krimigis and Sarris, 1980; Williams, 1981). Several explanations for the occurrence of these particles have been forwarded (e.g., Pellinen and Heikkila, 1978; Stern, 1979; Sonnerup, 1979b; Lyons and Speiser, 1982), none of which in our view, however, is as yet completely convincing (see also Vasyliunas, 1980). Aside from their intrinsic interest, high energy particles are also important as field line topology tracers but, as far as stress balance or energy density considerations are concerned, particles above a few tens of keV energy are typically quite negligible in the geomagnetic tail.

## V. THE DIFFUSION REGION

(A. B. Hassam)

One of the problems in applying the reconnection model to the Earth's magnetosphere is properly describing the diffusion region in which the term  $\mathbf{v} \times \mathbf{B}$  in Ohm's law is small. What then can balance the electric field in the diffusion region? The resistivity  $\eta$  associated with Coulomb collisions between particles is very small in space; the plasma is essentially collisionless. One possibility is the occurrence of anomalous resistivity in the diffusion region. In such a model, particles scatter off collective electric fields generated by a plasma microinstability, and this scattering process decouples the plasma from the magnetic field. Another possibility is that

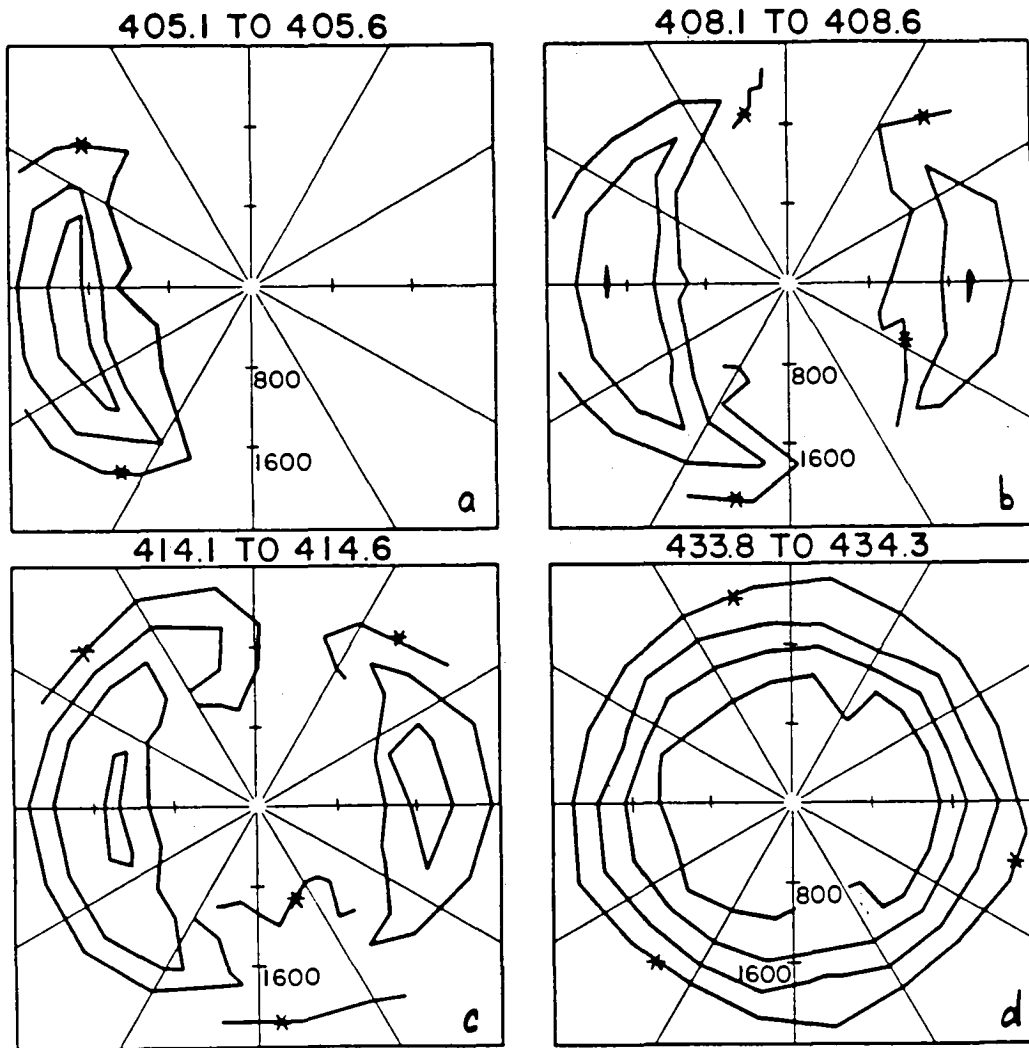


Figure 1-10. Observations of field-aligned accelerated ion beams observed on the outer boundary of the plasma sheet during substorm recovery. Each panel shows  $\log_{10}$  proton distribution function ( $f$ ) contours obtained from a 2D analyzer on ISEE 2 from 04:05 to 04:34 on March 1, 1978. The coordinates are solar ecliptic  $V_x$  and  $V_y$ , with the Sun and Earth to the left ( $+V_x$ ) and the dusk direction at the bottom ( $+V_y$ ). The contours with asterisks correspond to an  $f$  value of  $10^{-27.75} \text{ sec}^3 \text{ cm}^{-6}$  and the  $\log f$  values of adjacent contours increase by steps of 0.5. The average magnetic field for this period is aligned within  $1^\circ$  of the  $+V_x$  axis. In the top left hand panel a field-aligned earthward flowing ion beam is observed of characteristics shape. In the subsequent panels beam ions mirrored at Earth and moving tailward appear and the distribution develops towards isotropy within the plasma sheet proper (lower right hand panel). This is essentially the sequence for the penetration of the plasma sheet boundary layer described theoretically by Cowley (1980). (After Forbes et al., 1981.)

other terms in a generalized Ohm's Law may provide an effective "collisionless resistivity": in many field reversal regions, the ion gyroorbits may be quite large, necessitating a kinetic derivation of Ohm's Law.

A variety of microinstabilities have been studied over the past several years for the purpose of evaluating the anomalous transport needed for collisionless reconnection. A review of these instabilities and their role in reconnection processes has been given by Papadopoulos (1979) who argues that the analogy between classical and anomalous transport processes cannot be taken "too literally." The macroscopic evolution of the plasma in the presence of these different transport mechanisms can, in fact, be quite different. Thus, detailed theories and numerical simulations of collisionless reconnection processes are needed which self-consistently incorporate plasma microturbulence, rather than studies which extrapolate results from the collisional to the collisionless regime.

Two instabilities have been treated in detail in the literature with regard to collisionless reconnection: the ion acoustic instability (Coroniti and Eviatar, 1977) and the lower-hybrid-drift instability (Huba et al., 1977). The former is a kinetic instability which is excited via an electron-ion acoustic wave resonance when a sufficiently large relative drift between the ions and electrons exists in the plasma. For typical plasma conditions in the magnetotail where  $T_e \ll T_i$ , the inequality  $V_d > V_{cr}$  is required for instability. Here  $V_d$  is the relative drift velocity and  $V_{cr} \approx (kT_i/m_e)^{1/2}$  is the critical drift velocity. In order for a reversed field plasma to support this drift velocity, a rather thin current layer is required, viz.,  $\delta \sim c/\omega_{pe}$  where  $\omega_{pe}$  is the electron plasma frequency. Again, for typical magnetotail parameters ( $n \sim 1 \text{ cm}^{-3}$ ), this translates into an extremely small width  $\delta$  of the diffusion region, i.e., of the plasma sheet:  $\delta \lesssim 5 \text{ km}$ .

The lower-hybrid-drift instability does not require as strong a current as the ion acoustic instability. In the weak drift regime  $[(m_e/m_i)^{1/4} < V_d/V_i < 1]$ , where  $V_i$  is the ion thermal velocity, the instability is kinetic and is driven via a resonance between ions and a drift wave; in the strong drift regime  $[V_d/V_i > 1]$ , the instability is fluid-like and is driven via a coupling of a drift wave to a lower hybrid wave. For typical magnetotail parameters, the thickness of the plasma sheet necessary to excite this instability is  $\delta \sim 6r_{Li} \sim 2000 \text{ km}$  where  $r_{Li}$  is the mean ion Larmor radius in the lobe field (we have taken  $T_i = 4 \text{ keV}$  and  $B = 30 \gamma$ ).

A nonlocal theory of the lower-hybrid-drift instability in a reversed field plasma has been developed (Huba et al., 1980) to determine the structure of the perturbed wave field in the reversal region. A nonlocal theory is essential as most theories assume magnetized electrons, which is not the case in the vicinity of the neutral line. It was found (see Figure 1-11) that the waves extend over a substantial portion of the field reversed plasma but that they are strongest in the region away from the neutral line. In finite electron temperature plasmas, the waves are strongly damped in the vicinity of the neutral line. On the one hand, this localization away from the neutral line makes the nonlocal theory self-consistent with the magnetized electron assumption. On the other hand, it suggests that initial turbulence, in the absence of a quasi-linear profile modification, would not penetrate to the neutral line, precisely where the anomalous resistivity is desired.

The evolution of the magnetic field in the presence of a resistivity based upon the nonlocal mode structure of the lower-hybrid-drift instability was subsequently studied (Drake et al., 1981) and a transport equation for the magnetic field was developed for an arbitrary resistivity profile in a field reversed plasma. A model was used such that the resistivity  $\eta \propto B^2$ , i.e.,  $\eta = 0$  at the neutral line, but  $\eta \neq 0$  away from it. It was found that magnetic flux is transported towards the neutral line (see Figure 1-12b) and that the current density increases there (Figure 1-12a). This leads to the possibility that waves can subsequently penetrate the neutral region during the nonlinear evolution of the field reversed plasma. Such an evolution has been observed in particle

simulations of field reversed plasmas (Winske, 1981; Tanaka and Sato, 1981). However, more recent simulations using realistic mass ratios do not show any gross plasma or magnetic field changes due to the instability, other than strong electron heating. However, to observe significant profile modifications one would need simulation runs on time scales of quasi-linear diffusion which are considerably longer than those available to date.

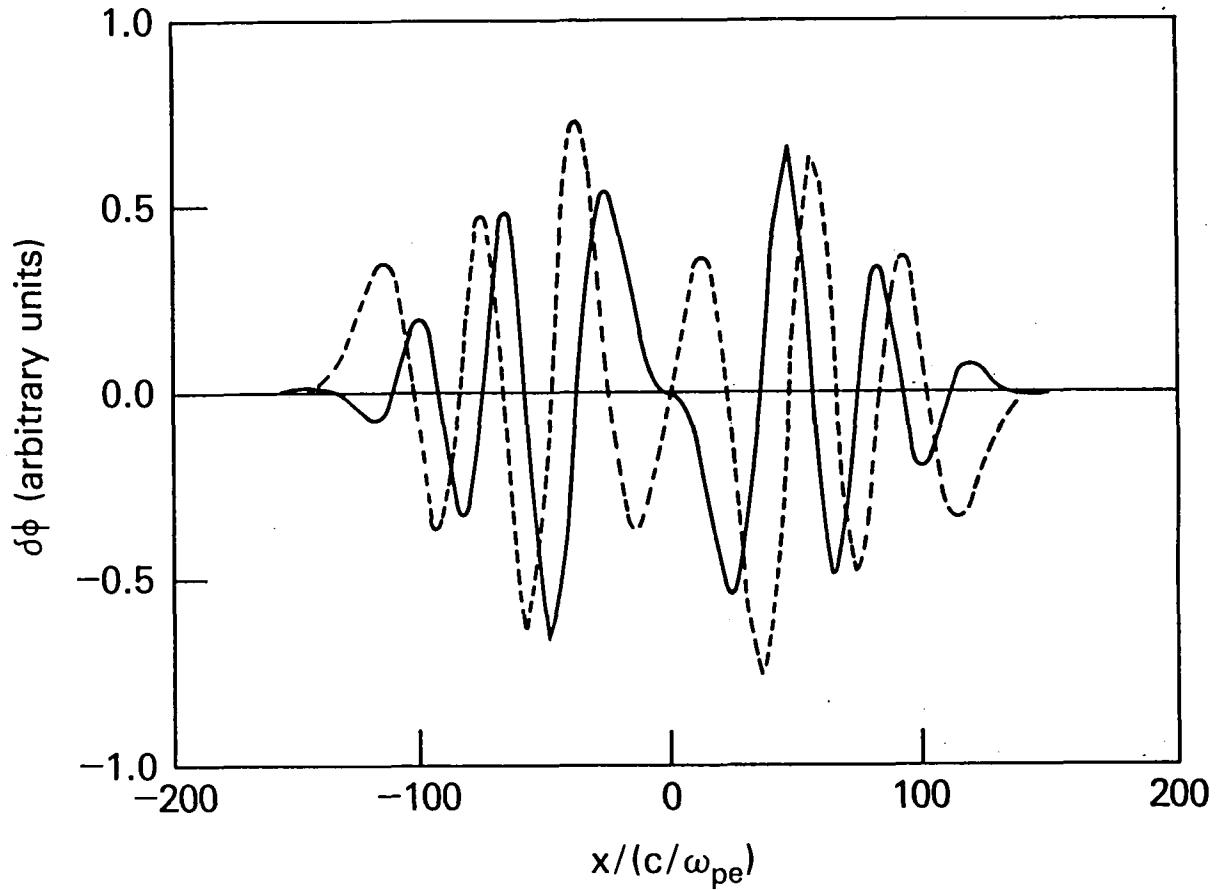


Figure 1-11. Mode structure of the lower-hybrid-drift instability in a field reversed plasma with cold electrons ( $B$  reverses direction at  $x=0$ ). The fluctuating potential  $\delta\phi$  is decomposed into its real (solid) and imaginary (dashed) parts.

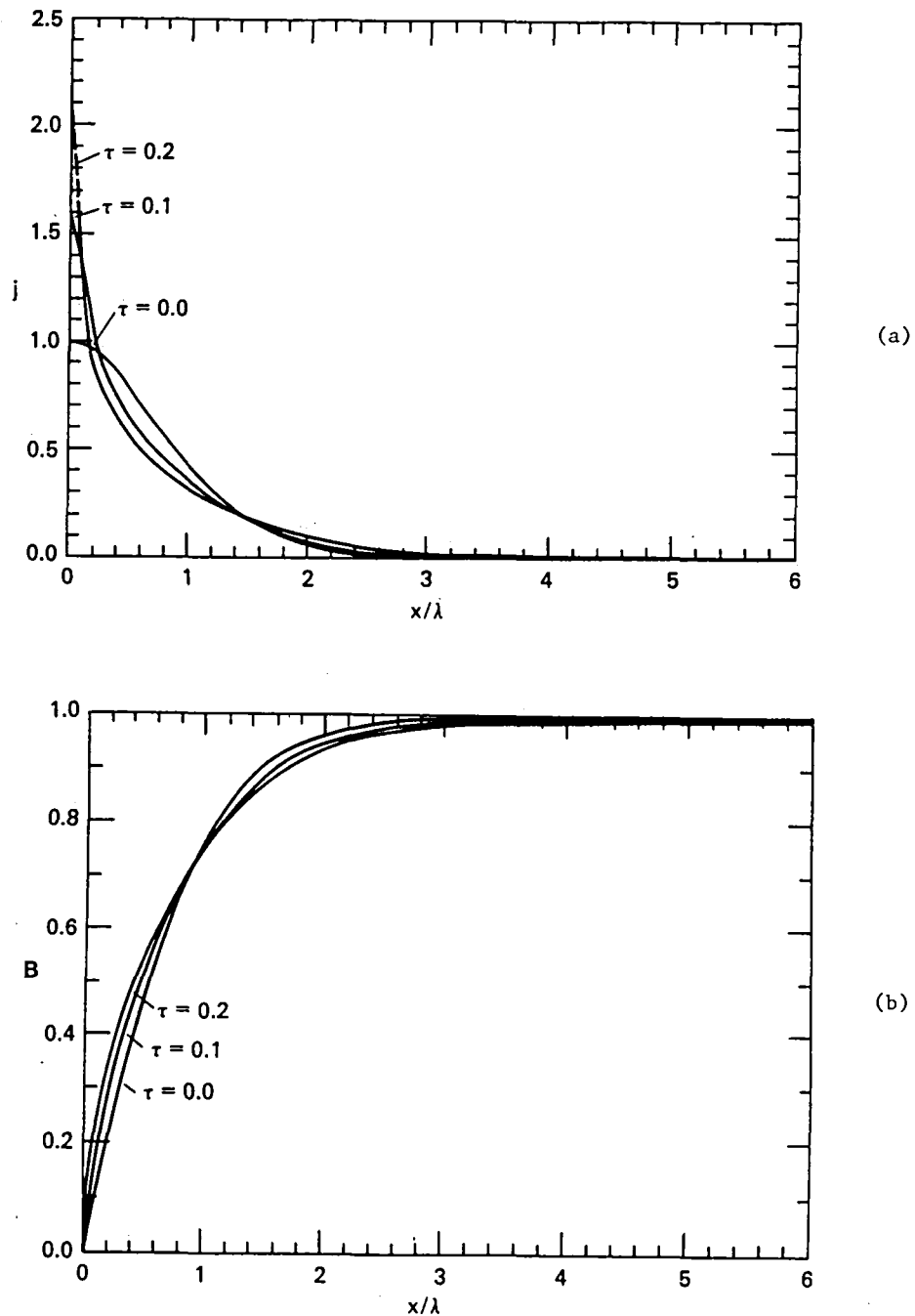


Figure 1-12. Time evolution of (a) the current density  $j$  and (b) the magnetic field  $B$  in the presence of a model anomalous resistivity ( $\eta \propto B^2$ ) based on the nonlocal mode structure of the lower-hybrid-drift instability. The symbol  $\tau$  denotes various times elapsed from the initial profiles at  $\tau=0$ . Note that magnetic flux is transported toward the neutral line and that the current density increases there.



The other possibility for reconnection, namely the direct onset, beyond some threshold, of a large scale collisionless tearing instability, must also be considered. In this case the frozen-in condition  $\mathbf{E} + \mathbf{V} \times \mathbf{B} = 0$  is violated when the particle inertia leads to non-adiabatic behavior in the neutral sheet. The plasma sheet would become unstable when the gyroperiod of ions based on  $B_z$ , exceeded the growth time of the tearing mode (Schindler, 1974; Galeev and Zelenyi, 1976). However, the ion-tearing mode might be stabilized by charge separation fields between the non-adiabatic ions and the still adiabatic electrons (Pellat, 1979). A nonlinear explosive mode has been proposed to overcome this apparent linear stabilization (Galeev et al., 1978).

A more rigorous evaluation of a two-dimensional energy principle (Goldstein and Schindler, 1982) has shown that a sufficiently stretched magnetotail can become unstable to a collisionless tearing mode even if the electrons are adiabatic. (In one case where instability occurred, the conditions for instability were  $r_{Li}/\delta < .35$  and  $B_z/B_0 < 1/45$ , where  $Z$  is in the direction normal to the current sheet and  $B_0$  is the lobe field.) The scale of the unstable mode in the tailward direction was found to be comparable to the equilibrium scale length in that direction. For this stability analysis, it was important to start from a realistic self-consistent equilibrium including the tail lobes where  $B_z$  changes sign and becomes negative as it can be proved in a general way that configurations with just one sign of  $B_z$  everywhere are stable (Birn et al., 1975).

## DIRECTIONS FOR FUTURE WORK

Although the processes in the diffusion region discussed here appear promising in some respects, investigation in this area is far from complete.

(1) We have only discussed two microinstabilities which initially appear promising in providing anomalous transport for collisionless reconnection: the ion-acoustic instability and the lower-hybrid drift instability. Other instabilities have been suggested but have not been dealt with in detail (Papadopoulos, 1979; Lee, 1982). It is clear that under "typical" magnetotail conditions neither of these instabilities is likely to allow reconnection to occur: the plasma sheet is simply too broad. Thus, "atypical" conditions are required, viz., a very thin plasma sheet or current layer ( $\delta < \text{a few hundred km}$ ), if these instabilities are to play a role. Whether or not such "thin" plasma sheets exist, at least locally in the magnetotail, will have to be determined by observations.

(2) The plasma and magnetic field configurations, such as the Harris equilibrium or the Petschek model, used to study the influence of instabilities on reconnection processes have probably been oversimplified. Other geometries should also be investigated. For example, recently Ugai (1983) has performed a forced reconnection simulation in which the density becomes somewhat rarefied at the neutral line, which in turn leads to a much larger drift velocity at the neutral line than would have been expected. Also, the oversimplified configurations have related the cross-field current to the plasma inhomogeneity (i.e., the diamagnetic drifts). It may be that the current system is more complicated and is dictated by "external" conditions, such as the solar wind-magnetosphere dynamo system. And, with the exception of the work by Pu et al. (1981) only Maxwellian distributions have been considered thus far which may be inadequate. Finally, in order to improve the theoretical modeling it would be useful for the theorists to interact more closely with space and laboratory experimentalists and with numerical simulation workers to determine the appropriate plasma and field configuration. A case in point is reconnection at the magnetopause. Here, the magnetic field usually has rotational character ( $B_y \neq 0$ ) and may be nearly force free. These features may suppress, for example, the lower-hybrid drift instability.

(3) Most of the theoretical work on anomalous transport has focused on electrostatic turbulence. However, since the neutral line region is a high  $\beta$  plasma, it is likely that any turbulence

generated in this region will be primarily electromagnetic [this was found in the particle simulations of Winske (1981) and Tanaka and Sato (1981); see also Section XI] so that an effort should be made to understand anomalous transport associated with electromagnetic turbulence.

(4) Detailed numerical simulations of collisionless reconnection processes are needed which self-consistently incorporate microturbulence (Papadopoulos, 1979; Coroniti and Eviatar, 1977). In this regard, different instabilities may arise in different parts of the diffusion region. For example, the lower-hybrid drift instability may dominate away from the neutral line while the ion acoustic instability may dominate in the immediate vicinity of that line. Two dimensional hybrid simulations (particle ions and fluid electrons) would be of use in studying such effects.

(5) More work is needed in collisionless tearing mode theory. For the ion-tearing mode, for example, the onset of instability would have to be investigated more in detail and a comparison made with observations.

## VI. LABORATORY SIMULATION OF MAGNETOSPHERIC RECONNECTION

(P. J. Baum)

Presently, laboratory terrella studies of magnetospheric reconnection are being carried out by Podgorny et al. (1978, 1980; called "PDP") (see also Dubinin et al., 1980) at the Space Research Institute in Moscow (USSR), by Baum and Bratenahl (1982a; called "BB") in the T-1 device (see Figure 1-13) at the University of California, Riverside (USA), and by Minami and Takeya (1983; called "MTT") at Osaka City University (Japan). Each experiment is designed to satisfy a principle of limited simulation which allows a certain group of events to be modeled within the appropriate plasma regime. Each fulfills the following prescription: the magnetopause standoff distance considerably exceeds the terrella size; the Alfvén and sonic Mach numbers are much larger than unity; the solar wind duration is many magnetosphere transit times long; the magnetic Reynolds number divided by the product of the ion gyrofrequency and the ion collision time is much larger than unity; the magnetopause standoff distance is more than one gyroradius; and the ion collisional mean free path exceeds the magnetopause standoff distance. On the basis of this last condition, the MTT experiment seems semi-collisional.

### EXPERIMENTAL RESULTS

Using magnetic probe measurements compiled over many shots, PDP demonstrated that a Dungey-type open magnetosphere was obtained for an Alfvén Mach number of 1.5. This was subsequently verified by BB in the Alfvén Mach number range of 1.0 to 2.5. Figure 1-14 shows the T-1 magnetic map at Alfvén Mach number 2.5. Recent T-1 photographs, taken in the light from visible tracers stimulated by auroral particle precipitation, show structure in agreement with the magnetic map earthward of the magnetopause and sunward of the X-point in the tail. They also indicate an interesting east-west asymmetry in solar wind plasma entry through an elongated rather than a funnel-shaped polar cleft structure.

Using a powered double probe to inject luminous plasma along field lines, MTT find the nightside magnetosphere to be magnetically interconnected with the solar wind. Significant changes in the amount of interconnection occur in less than ten microseconds while the ohmic diffusion time is in the millisecond range. Neither PDP nor MTT comment explicitly on the time scale for collisionless reconnection although PDP may have implied that reconnection occurs because of plasma microturbulence or anomalous resistivity. Baum and Bratenahl (1982b) suggest that the Hall or inertial resistivity plays a large role in determining this time scale.

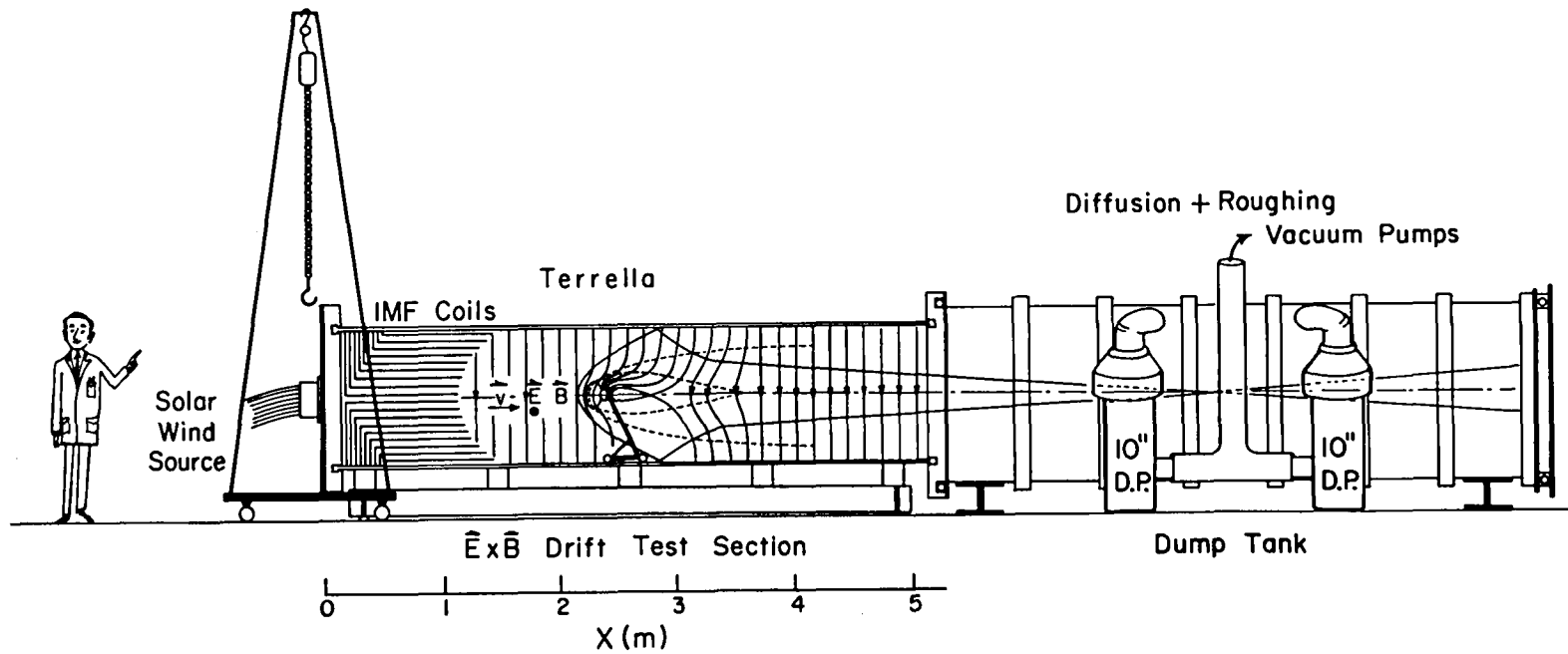


Figure 1-13. Schematic drawing of the UCR T-1 eleven meter terrella device. A coaxial plasma accelerator produces an ionized hydrogen stream of actual solar wind speed which interacts with an external IMF and model earth dipole field. The solar wind duration is 100 microseconds. (From Baum and Bratenahl, 1982a.)

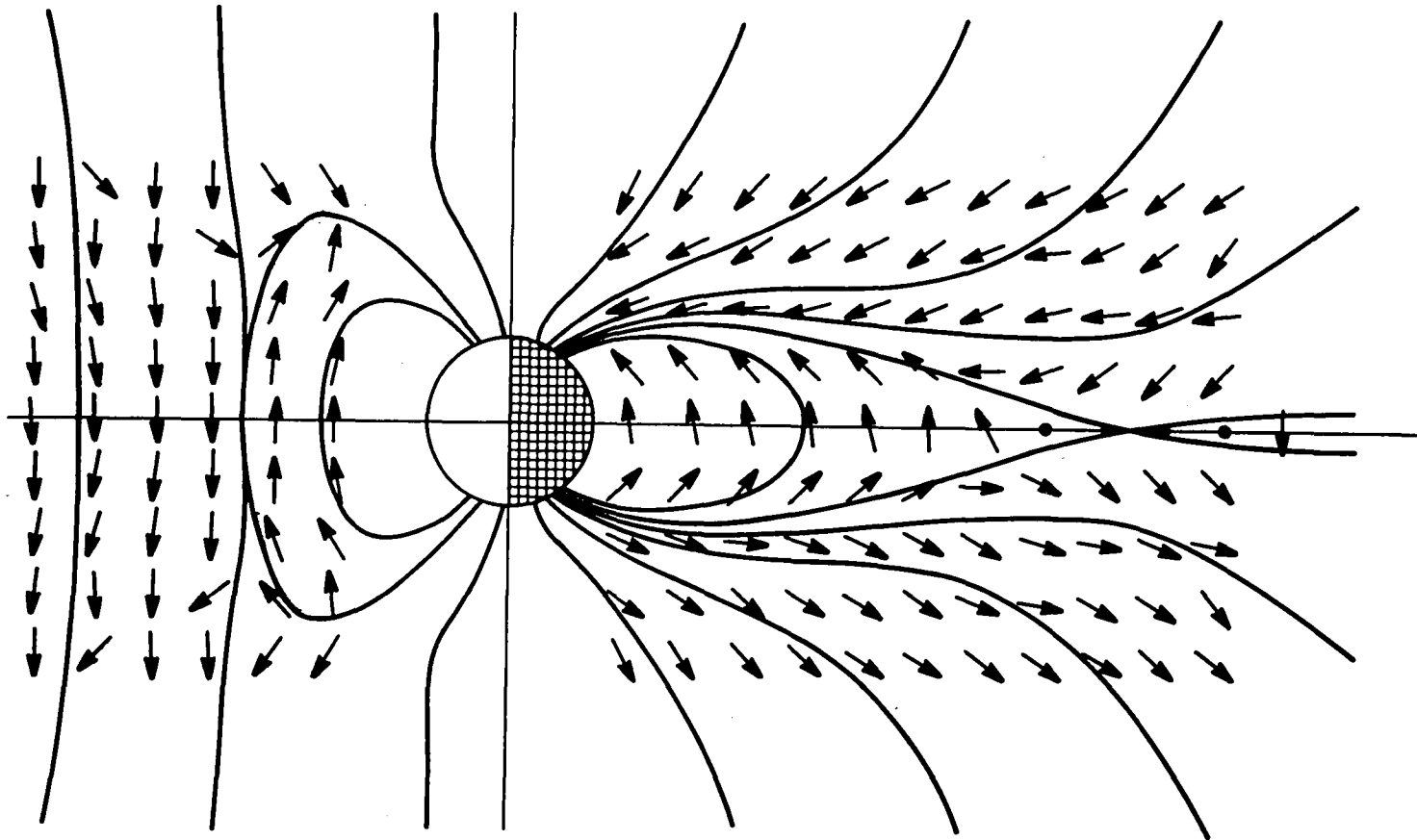


Figure 1-14. Typical magnetic field map of the UCR T-1 magnetosphere for a southward IMF at Alfvén Mach number 2.5 during the fast plasma phase. (From Baum and Bratenahl, 1982a.)

For Alfvén Mach numbers much larger than unity, the plasma kinetic energy density greatly exceeds the magnetic energy density. From a total energy point of view, a solar wind at Alfvén Mach number 10 differs only one percent from a Chapman-Ferraro solar wind with an Alfvén Mach number of infinity. Thus closed models are a good zero-order approximation to the real magnetosphere in certain respects such as the magnetosphere size and shape. However, the differing effects on magnetic topology, polarization, and convection at Alfvén Mach numbers of ten and infinity are significant as discussed below.

PDP also performed studies at an Alfvén Mach number of ten finding that a large magnetic island (visor) of thickness comparable to an ion gyroradius formed at the subsolar point and that neutral points formed near the polar clefts as shown in Figure 1-15. Sonnerup (1979a) suggested that the island should be one ion gyroradius thick and that it might be subject to patchy reconnection. Using faster data acquisition equipment and increasing the length of the large island to 0.4 meters by the use of a larger terrella, PDP then observed the formation of a chain of smaller islets within the main island. They found that the characteristic size of the smaller islets was 0.2 to 0.3 ion gyroradii or 2-4 cm and interpreted this structure in terms of the tearing mode instability. At the equator there was an equal probability of observing either an X-type neutral point or an O-type neutral point. However, near the polar clefts the field was always consistent with the visor geometry. The islets shifted according to increasing or decreasing solar wind ram pressure resulting in a change from an initial O-point to an X-point at the subsolar point at a later time.

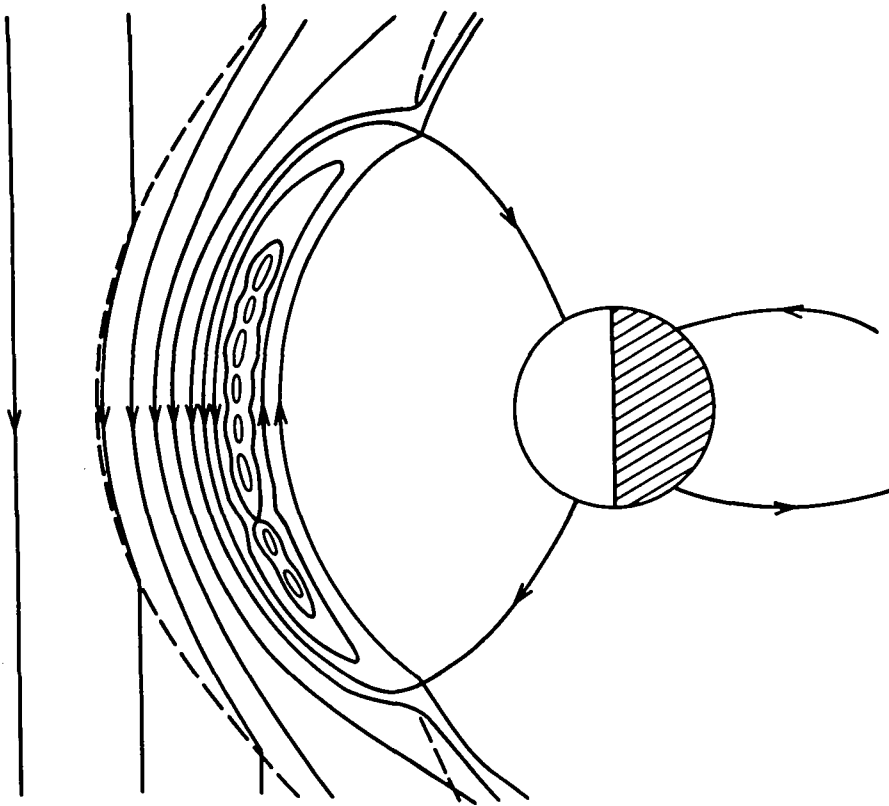


Figure 1-15. Pattern of the magnetic field distribution on the dayside of the magnetosphere with a southward IMF for Alfvén Mach number 10. Reconnection in the polar clefts leads to the formation of a large magnetic island in front of the magnetosphere. The interior of the island tears into chains of smaller islets. (Adapted from Dubinin et al., 1980.)

PDP had thus established that even for a purely southward IMF the Dungey model did not match their observations at high Alfvén Mach number. PDP found that reconnection occurred near the cusps while Dungey predicted that it would occur at equatorial height. The case where the IMF is not purely southward is equally interesting. Crooker (1979) had published a qualitative model in which reconnection occurs in those regions where the IMF is antiparallel to the lines of a closed magnetosphere. This led to the prediction that with an east-west interplanetary magnetic field (IMF) component, the separator line would lose the degeneracy which it had in the Dungey model by splitting into two branches which link to the cusps. Thus reconnection would primarily occur at the cusps. PDP found that cusp reconnection again occurred with an east-west IMF component added and the island structure also persisted. PDP reported that with an increasing east-west component the island's magnetic field increased in helicity while an interconnection developed between the IMF and island's field. It remains to be clarified how the IMF can consistently interconnect with the island itself. Thus in spite of remarkable advances in experimental and theoretical approaches, further surprises seem likely in understanding the general three-dimensional structure of the magnetosphere.

PDP reported that, with a southward IMF, the sign of the dawn-dusk electric field above the north polar cap corresponded to convection away from the Sun. UCR T-1 results confirm this point. With a northward IMF the convection was toward the Sun and with no IMF there was still significant antisunward convection. PDP interpret these results as evidence for two simultaneous convection mechanisms: reconnection and viscous drag. We may conclude that the magnetosphere is electrically polarized without any IMF but becomes more strongly polarized when interconnected with a southward IMF. PDP found that while the polar cap electric field reverses sign with a reversal of the IMF, the polar cleft electric field measured 3.5 cm above the terrella along a dawn-dusk line, was unidirectional, and corresponded to convection away from the Sun.

## DISCUSSION

In the foregoing experiments, as well as in the real magnetosphere, greater attention should be given to the time scale of collisionless reconnection. Experimentally, reconnection is observed to occur rapidly in the absence of interparticle collisions where the ordinary conductivity is very large. One could be content with the invocation of current-driven wave-particle turbulence to explain the rapid time scale but for the fact that sufficiently large current densities are often not observed (see Section V). Collisionless reconnection theory concentrates on flow parameters such as the Alfvén Mach number and on structural length scales without addressing the time scales involved in the more general problem. Vasylunas (1975) discussed some inertial effects while Sonnerup (1979b) also included the perception that Hall effects may be prominent in the "diffusion" region of the steady reconnection problem. Hall effects arise from the differing inertia of ions and electrons in a magnetic field.

The transient initial reconnection process (the phase where the field lines become interconnected) requires that the line integral of the inductive electric field along the separator be nonzero. The ohmic electric field is inadequate for this purpose. From a two-fluid point of view we must examine the other terms in the generalized Ohm's law such as the Hall, electron inertial, and pressure tensor electric fields. When this is done it is found that many of these terms are of the requisite magnitude for agreement with experiment in the absence of turbulence. Thus it appears that there is a need to incorporate the entire generalized Ohm's law into reconnection simulations.

The resistive diffusion time is defined by  $t = \mu_0 \sigma \ell^2$  where  $\sigma$  is the conductivity and  $\ell$  is a characteristic length. If one employs the gyro (Speiser, 1970) or Hall (Baum and Bratenahl, 1982b) conductivity,  $\sigma = ne/B$ , one obtains the time estimate  $t = \mu_0 \ell^2 ne/B$  which is in reasonable

agreement with measurements under some conditions. Physically, this would happen if the current sheet carriers do not remain in the sheet indefinitely but are ejected after a time comparable with an ion gyroperiod as a result of particle motion in the complex electric and magnetic fields of the current sheet. There is a need to understand the relative contributions of inertial effects and wave-particle turbulence to the resistivity of collisionless plasmas so that critical observational tests can be developed for each.

Finally, a laboratory simulation of impulsive flux transfer events should be performed.

## VII. RECONNECTION AND THE EVOLUTION OF SOLAR FIELDS

(S. W. Kahler)

Until recently, observational examination of the role of magnetic reconnection both in the evolution of solar fields and in solar flares has been hampered by the fact that optical data generally provide a look at magnetic structures only in the photosphere or chromosphere, i.e., only at the thin slices of magnetic flux lines where they intersect those atmospheric layers. Solar coronal observations providing a three-dimensional view of magnetic fields were first provided by the X-ray and EUV instruments on Skylab. Coronal emission regions, observed with spatial scales of several arcsec ( $1 \text{ arcsec} = 725 \text{ km}$ ) and time scales of seconds to months, proved to be arches or loops of plasma confined by magnetic fields. These Skylab coronal images of magnetic arches have given us our best opportunity to observe solar magnetic reconnection.

Solar fields possess several properties which make their observation somewhat complementary to those of the geomagnetic field. In the low corona the low beta plasmas allow the magnetic fields to dominate the plasma, but in general the emission from plasmas on specific magnetic structures is bright enough to allow us to image those structures. Also the coronal scale heights are large compared to most loops so that uniform emission occurs along the lengths of the coronal loops. The result is that a synoptic view of entire coronal loops is obtained which is not achieved in geomagnetic observations. Some evidence exists that, at least in active regions, the plasma pressure, and hence brightness, scales with a power of the photospheric magnetic field of the loop (Golub et al., 1982). In principle we can also gauge any mass motions resulting from magnetic reconnection through X-ray or EUV emission line Doppler shifts, routinely measured to tens of  $\text{km s}^{-1}$  (Cheng et al., 1982). There are observational problems such as trying to untangle line-of-sight effects, that compromise the use of solar data for understanding magnetic reconnection, but the advantages described above give cause for optimism.

Although solar magnetic-field reconnection is best known in the context of flares, it is also suggested in the Babcock (1961) and Leighton (1964) model of the solar magnetic field cycle. In that model, bipolar fields emerge in active region latitudes ( $\sim 0$  to  $40^\circ$ ) with the preceding portions of the bipoles diffusing to the equator and the trailing portions eventually migrating to the solar poles. The polarities of the fields moving to the equator are opposite on the two sides of the equator, so that a buildup of flux around the equator is prevented only by annihilation of opposite fields. Similarly, the polar fields are reversed with each cycle as the polarities of the trailing portions of bipoles also reverse and cancel the existing polar fields through the presumed processes of poleward diffusion and magnetic field annihilation. In this section the role of reconnection in these evolutionary processes is discussed.

## BIPOLAR MAGNETIC REGIONS

Sheeley et al. (1975a, b) have presented XUV observations from the NRL Skylab spectroheliograph in which the emergence of a bipolar magnetic region (BMR) was followed within a day or two by the appearance of a loop connecting one of the BMR polarities to pre-existing background fields of opposite polarity. They suggested that the time scale of this process may keep pace with the rapid emergence of new magnetic flux in the photosphere, i.e., that they are short compared to those ( $\sim$  hours) over which one characteristically sees significant changes in the flux patterns. They speculated that reconnection as described by Petschek (1964) occurs as field lines are forced together by their moving footpoints.

Svestka et al. (1977) have interpreted as reconnection the X-ray appearance of a transequatorial loop connecting a new BMR in the north to an old BMR in the south (Figure 1-16). Unfortunately, the loop appeared after a 2 1/2 day gap in the observations, but there was no earlier indication of its existence. They suggested that loops interconnecting different active regions are formed in two basic ways. Loops crossing the equator are produced by reconnection of two loops extending toward each other from the two bipolar regions, while in most cases the emergence of the BMR simply causes a pre-existing loop with its footpoint near the BMR to brighten.

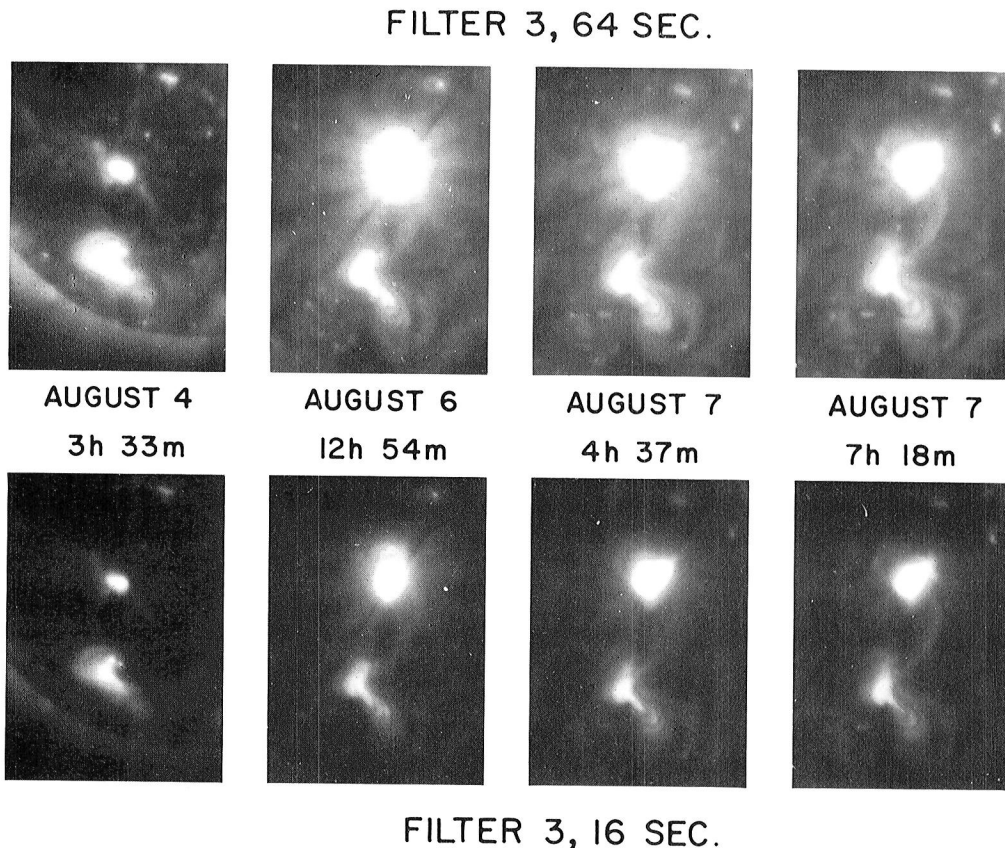


Figure 1-16. Possible birth by magnetic reconnection of a transequatorial loop connecting the old active region McMath 12472 (below) to the newly emerged active region McMath 12474 (above). Skylab X-ray images with two photographic exposure times in the 2-32 and 44-54 Å waveband are shown. (From Svestka et al., 1977.)



Nolte et al. (1977) have used coronal X-ray and XUV observations of loops to analyze large-scale reconfigurations, defined as the formation, disappearance, or motion of loops on a scale comparable to their size. They pointed out that a substantial transient brightening can occur in a pre-existing invisible magnetic loop, giving the appearance of the formation of a new loop and that not all field lines belonging to a large-scale configuration in the corona are visible as X-ray or XUV emitting structures. Besides reconfiguration of potential fields Nolte et al. give frozen-in motion of fields and plasma and reconnection as the other two possibilities for large scale magnetic field reconfiguration. For reconnection they argue that the field merging rate must be on the order of the velocity of the photospheric field motion, generally  $\geq 0.1 \text{ km s}^{-1}$ . Reconnection is assumed to result from current instabilities with high electron drift velocities,  $V_d$ . Letting the current be  $B/4\pi\delta = n_e q V_d$ , adequately high drift velocities are achieved in coronal conditions only for  $\delta \sim 5 \text{ m}$ . Thus, the predicted thickness of the diffusion region is extremely small in these applications. However, as we have seen (Figure 1-2b, for example), the outflow regions, which contain the energized plasmas, will be enormously larger than the diffusion region and may well be observable by space telescopes.

Nolte et al. argue that the use of additional information, such as photospheric magnetic field measurements, can be used to demonstrate whether reconnection has occurred. As an example, they show Skylab X-ray images of an emerging BMR which appears to result in large loop connections to the pre-existing surrounding fields. From the magnetic field data they infer that these loops would have to connect identical magnetic polarities in the BMR and surrounding fields. They conclude that in this case no reconnection has occurred and that the loops were structures pre-existing from before the emergence of the BMR. If this reasoning is applied to the example cited by Sheeley et al. (1975a), one concludes that reconnection did occur (although Nolte et al. question the conclusion that reconnection must have occurred).

## CORONAL HOLE BOUNDARIES

A solar coronal structure of perhaps greater relevance for magnetic reconnection than the BMR loops is that of the coronal hole boundary. Coronal holes, large regions nearly devoid of any X-ray emission, were intensively studied by Skylab investigators who interpreted them as due to magnetically open regions on the Sun (Figure 1-17). Maxson and Vaiana (1977) claim that the holes are completely open while the large scale emitting structures are entirely closed, but evidence to the contrary exists (Levine, 1982). The coronal holes make a nice magnetic analogy to the Earth's polar cap, and the question of which field lines are open and which closed is easily studied at hole boundaries where the difference in brightness between the two is usually pronounced. Nolte et al. (1978a, b), using Skylab X-ray images, divided all boundary changes into two groups: small scale and large scale. Small scale changes comprise about 70% of the cases. They could easily result from random motions of the underlying photospheric magnetic fields and thus have nothing to do with reconnection of coronal fields. The remaining 30% were large scale changes, i.e., changes over a boundary length greater than three times the average supergranulation cell length ( $\sim 9 \times 10^4 \text{ km}$ ). They accounted for most of the evolution of coronal hole areas and were interpreted as the result of field line reconnection. Although the time resolution of the X-ray data was never more than six hours, and in many cases was as short as tens of minutes, only images spaced about 24 hours apart from each other were examined, precluding any close examination of the time scales for the large scale changes. The best estimate of this scale comes from the Solodina et al. (1977) study of the birth of a coronal hole which was found to be associated with an upper limit of 9 1/2 hours for a 40% decrease in X-ray intensity. The coronal hole studies seem to show evidence of true field reconnection, but no attempt has been made to identify the size and time scale of discrete reconnection events.

## CORONAL HOLE 4

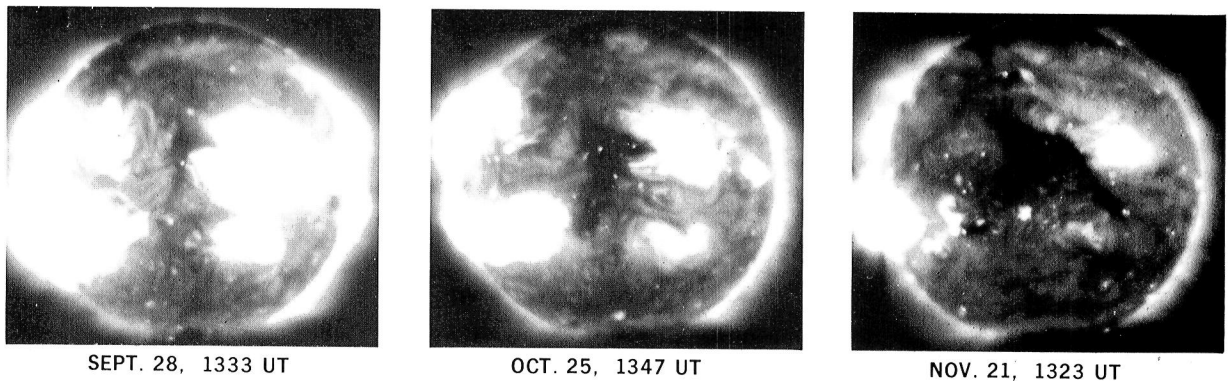


Figure 1-17. First three central meridian passages of Skylab coronal hole 4 shown as the dark region near disk center in these soft X-ray images. The holes are regions of open magnetic fields which grow or shrink due to presumed magnetic reconnection events at the boundaries. (From AS&E.)

## MAGNETIC FLUX DISAPPEARANCE

Magnetic flux continually emerges from below the photosphere in bipolar structures which then grow in size. The emergence of flux must be balanced, at least over long time scales, by a corresponding loss of flux. Howard and LaBonte (1981) examined 13 1/2 years of Mt. Wilson solar magnetograms and found that magnetic flux is highly concentrated toward the equator, with only about 1% of the total flux in the polar regions. The field annihilation is concentrated in the active latitudes, and about 10% of the total flux disappears each day, balanced by a corresponding flux emergence rate. An investigation of changes in the magnetic flux in active regions associated with the disappearance of sunspots (Wallenhorst and Howard, 1982) showed that an amount of flux greater than that observed in the spots themselves, vanished on or near the day of spot disappearance. The opposite-polarity portions of the regions were too far apart for any large-scale cancellation of fields, so it was suggested that flux bundles might sink beneath the surface. In comparing Skylab spectroheliograms and photospheric magnetic field maps, Sheeley et al. (1975b) also noted this flux loss from active regions. Recently, Wallenhorst and Topka (1982) used H $\alpha$  images and magnetograms to observe sunspot disappearance in an active region. They observed a 40% decrease in flux over the four-day interval. No evidence was found of fields diffusing away from the active region or of field lines moving toward the line of zero vertical field, which might be expected if flux loops were sinking into the photosphere. In summary,

there is good evidence that the magnetic flux in dying active regions decays at the rate of  $\sim 10\%$  per day and that this loss occurs in the active region itself. The lack of evidence for diffusion or flux tube sinking leaves reconnection as a strong possibility, but no direct evidence for it has yet been presented.

## DISCUSSION

While the emergence of magnetic flux on the Sun must be balanced by a corresponding removal of flux at the same average rate, the role of reconnection in this removal is still unknown. Processes other than reconnection, such as current dissipation, may well be important, so we cannot assume that loss of flux in a given case must be due to reconnection. An opportunity to measure the possible analog of geomagnetic flux-transfer events exists with the observations of coronal hole boundary changes. While such measurements will not confirm the existence of reconnection, they can at least put limits on the size and time scales of the presumed reconnective events.

## VIII. RECONNECTION IN SOLAR FLARES

(T. G. Forbes and E. R. Priest)

Detailed flare observations obtained over the past 15 years, especially those from Skylab (Sturrock, 1980) and the Solar Maximum Mission, suggest that there are two basic types of flare, namely simple-loop (or compact) flares and two-ribbon flares. Most flares are small and of the first type in which a single loop brightens rapidly and then slowly fades, preserving its shape and position in the process. It is quite possible that these small, simple-loop flares involve magnetic reconnection either through the onset of the kind of resistive kink instability that occurs in tokamaks (Coppi et al., 1976; Sykes and Wesson, 1976; Spicer 1977; Schnack and Killeen, 1979; see also Section IX) or through the emergence of magnetic flux loops into the corona from the photosphere (Heyvaerts et al., 1977). However, it appears equally likely at the moment that they do not involve magnetic reconnection directly but are instead the result of either an ideal MHD kink instability or a thermal instability (Giachetti et al., 1977; Hood and Priest, 1979, 1981; Van Hoven, 1981; Mok and Van Hoven, 1982).

On the other hand, the two-ribbon flare, the second of the two basic types, seems almost certain to involve a significant amount of magnetic reconnection. These flares occur inside complex active regions and involve extensive changes in the magnetic field configuration during the course of their evolution. All large events ( $\geq 10^{32}$  ergs) are of this type, and their scenario is as follows: a dense filament supported in or above an arcade of coronal magnetic loops erupts outward, away from the Sun, at the start of the flare. As the eruption proceeds, two parallel ribbons of emission appear down in the chromosphere. These ribbons lie at the footpoints of an overlying arcade of hot ( $> 10^7$  K) coronal plasma loops which form at the same time. During the main phase of the flare (lasting from 1-30 hours) the ribbons continually move apart while the hot loops suspended over them slowly rise at a speed which is typically  $1-20 \text{ km s}^{-1}$  (for further discussion see Priest, 1981, 1982).

The generally accepted explanation for the above sequence of events is that the initial, pre-existing filament acts as a tracer of the motion of a magnetic arcade which erupts due to the onset of a magnetic instability (e.g., Van Tend and Kuperus, 1978; Hood and Priest, 1980; Birn and Schindler, 1981). This creates an open field, which then closes back down to a lower state of magnetic energy (Kopp and Pneuman, 1976; Pneuman, 1981). The final stage, when the field is reconnecting back down, is shown in Figure 1-18. Here one sees the open magnetic field lines

moving toward a magnetic neutral line (i.e., an X-line or separator) where they reconnect to form a set of closed magnetic loops below the neutral line. As the open field lines continue to reconnect, the region of closed magnetic loops increases and the neutral line moves upward. The large amount of magnetic energy which is released by this reconnection process is mostly converted into hot ( $10^7$ - $10^8$ K) plasma by both slow and fast MHD shocks existing in the region just below the neutral line and above the closed loops (Cargill and Priest, 1982; Forbes and Priest, 1983). As this hot plasma is radiatively cooled, it flows downward along the magnetic field, thereby forming the cool plasma loops and the emission ribbons.

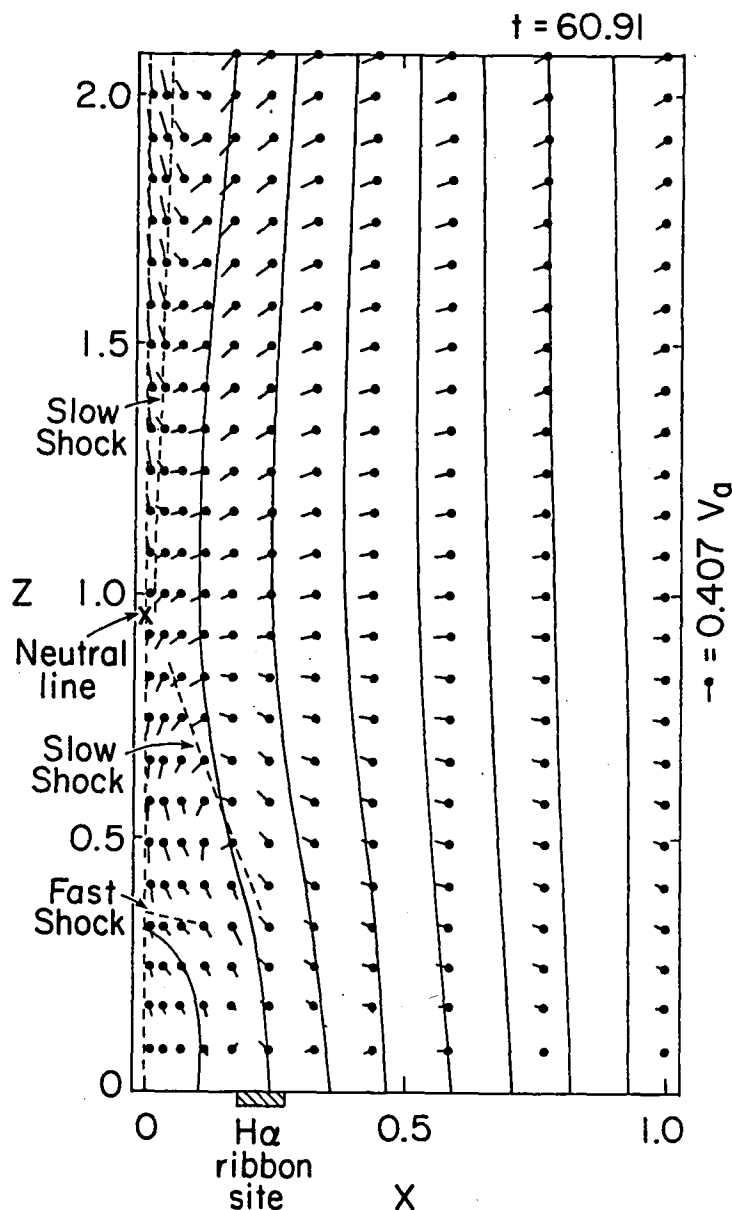


Figure 1-18. Magnetic field lines and flow vectors during the numerical experiment of line-tied reconnection, indicating positions of the slow and fast shocks (Forbes and Priest, 1982a, 1983).

Recently, a numerical simulation of this line-tied reconnection process has been reported by Forbes and Priest (1982a, 1983; see also Figure 1-18). They start with an open vertical magnetic field, tied at the base, and follow the growth of reconnection through a linear tearing phase to a nonlinear phase of Petschek reconnection with slow shocks and an inflow speed of 0.06 times the external Alfvén speed. The external plasma beta is 0.1 and the magnetic Reynolds number is  $10^3$ . The reconnection point rises at up to 0.12 times the inflow Alfvén speed. An interesting feature is the appearance of a fast-mode shock wave below the neutral point where the jet of plasma emitted from the reconnection site is rapidly slowed down (see Yang and Sonnerup, 1976). Also, the current sheet may become stretched out so much that the Petschek mode itself becomes unstable to the creation of secondary neutral points by the tearing mode. This creates new magnetic islands which subsequently undergo rapid annihilation by coalescence instability. The resulting intermittent energy release may account for oscillations that are observed in X-ray and radio flux. However, it is difficult to extrapolate confidently to the higher magnetic Reynolds numbers appropriate to the Sun, where the reconnection may well be more turbulent in nature (Matthaeus and Montgomery, 1981; see Section X).

With the aid of X-ray observations from Skylab, Moore et al. (1980) have been able to track fairly accurately the locations of the hot plasma loops and the emission ribbons almost continuously throughout the lifetime of these features. From these positional measurements one can then obtain the speed at which the hot loops and ribbons are moving. From the point of view of reconnection theory these measured speeds are of special importance because they are directly related to the rate at which the reconnection is occurring in the flare. This is because the motions of the ribbons and the loops are a direct result of the topological changes associated with the reconnection and are not actually the motion of the plasma itself. One can show from Faraday's law that the electric field,  $E_O$ , at the neutral line is related to the ribbon speed,  $\dot{X}_O$ , by

$$E_O = \dot{X}_O B_z(X_O, 0)$$

where  $B_z(X_O, 0)$  is the radial component of the magnetic field at the location,  $X_O$ , of the ribbon and at time  $t=0$  (Forbes and Priest, 1982b). Therefore observing the radial component of the chromospheric magnetic field and the speed of the ribbons gives a direct measure of the reconnection electric field,  $E_O$ , along the X-line. Similarly one can also relate the apparent motion of the hot plasma loops to  $E_O$ , though this involves a somewhat more complicated functional dependence upon the coronal magnetic field below the loops.

Analysis of the event shown in Figure 1-19 indicates that the reconnection rate, as measured by  $E_O$ , was nearly constant throughout the time period when the hot loops and associated ribbons were visible. This suggests that the reconnection was occurring at the maximum allowable rate for almost the entire event. For Petschek reconnection one expects

$$E_O = B_C A_C M_A$$

where  $B_C$  is the coronal magnetic field,  $A_C$  is the coronal Alfvén velocity and  $M_A$  is the Alfvénic Mach number of the plasma flowing in towards the neutral line when this plasma is still located well outside the reconnection region. In the compressible version of Petschek's theory (Soward, 1982; Soward and Priest, 1982) the maximum allowable value of  $M_A$  is

$$M_A = [\pi\gamma/(4\gamma-2)] [\lambda\eta(M_A R_m^2)]^{-1}$$

where  $\gamma$  is the ratio of specific heats in the plasma. Also,  $R_m$  is the magnetic Reynolds number defined by  $A_C L/\eta$ , where  $L$  is the dimension of the hot plasma loops and  $\eta$  is the electrical resistivity. Although no attempt has yet been made actually to determine  $E_O$  for a set of events, one can estimate  $E_O$  and  $M_A$  on the basis of the values of  $X_O$  and  $B_C$  typical of two-ribbon

flares. The estimate so obtained implies that  $R_m$  in the main phase of the two-ribbon flares is  $\approx 10^8$ . Since the order of magnitude of this value is only slightly smaller than that of the value of  $10^{10}$  predicted by classical plasma transport theory for the corona, there appears to be no need to invoke extensive anomalous resistivity for the main phase of a two-ribbon flare. More observations and analyses are needed in order to shed light on this problem, and there is a need to study not only large scale events but also the possible numerous micro-flarings and associated magnetic field reconnections that may be responsible for keeping the corona at its temperature of  $2 \times 10^6 \text{K}$  (Heyvaerts and Priest, 1983; see also Section X).

The arcade eruption may be caused by a tearing mode instead of an ideal instability. Steinolfson and Van Hoven (1983a; 1983b) have recently analyzed the tearing mode for values of the Lundquist number,  $S$ , and wavelength that are more appropriate to solar conditions than hitherto. At long wavelength, the constant- $\psi$  approximation fails and the mode grows linearly at a rate proportional to  $S^{2/3}$  with a large region affected by it. In the nonlinear regime, growth is much faster than that of the Rutherford regime and the maximum electric field is only one millionth of the Dreicer value. Furthermore, in a series of important papers Steinolfson (1983), Steinolfson and Van Hoven (1984), and Van Hoven et al. (1983) have incorporated an energy balance that is realistic for the solar corona. For solar  $S$  values they find the normal tearing mode and a radiative mode that is 100 times faster and involves a substantial amount of a reconnection and energy release.

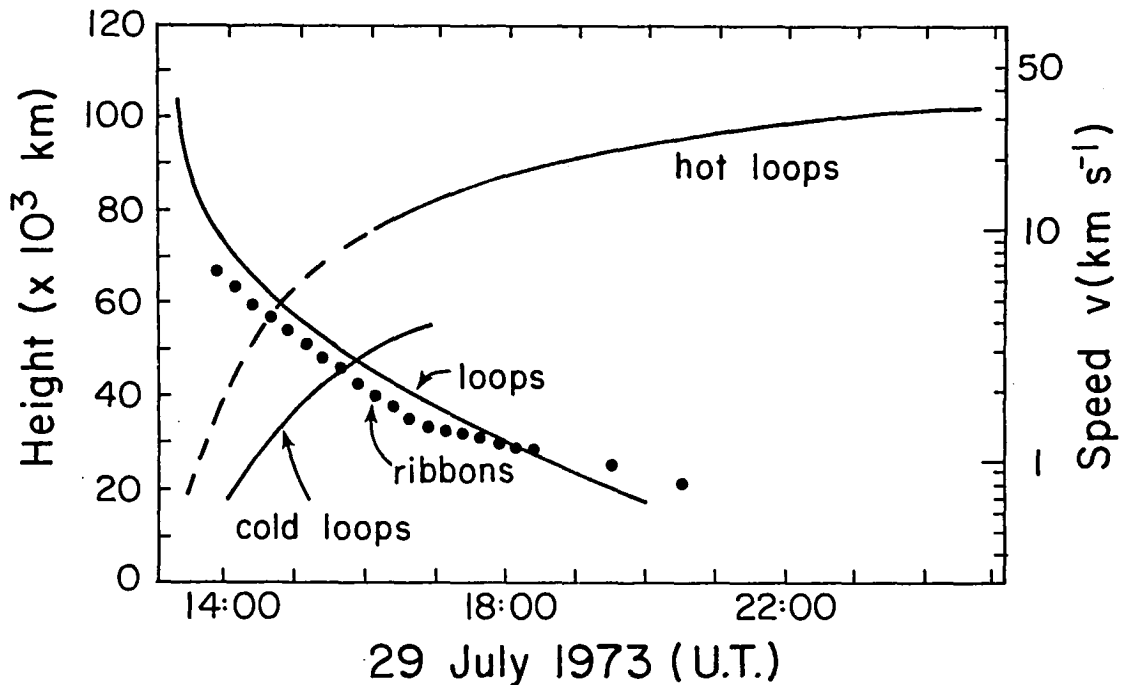


Figure 1-19. Observations from Skylab of the height and velocity of the loops below the rising neutral line as functions of time during a major flare (after Moore et al., 1980).

## SUMMARY

The conclusion that magnetic field reconnection in some form plays an important role in solar flares appears unavoidable, but the detailed manner in which the process operates remains controversial. The presence of visibly emitting plasma in the solar atmosphere makes it possible to discern the global magnetic field structure throughout the evolution of a flare. This global type of observation is not ever likely to be feasible in the magnetosphere, but it is becoming increasingly important in the case of the solar atmosphere as previously unobtainable high spatial resolution measurements become available from Earth orbiting observatories. Such observations, in conjunction with numerical simulations, will allow us a unique opportunity to use the measurement of topological changes in solar flare events as a diagnostic for determining the reconnection rate in an astrophysical plasma.

## IX. ACCELERATION OF SOLAR-FLARE ELECTRONS BY RECONNECTION

(D. S. Spicer)

Reconnection is, without doubt, the most popular flare mechanism ever proposed. This process is attractive for two fundamental reasons: it allows magnetic energy stored globally to be locally converted into thermal energy and convective flows, while simultaneously giving rise to large induced electric fields, and thus particle acceleration, parallel to the magnetic field. Two basic types of flare reconnection models are found in the literature: (i) Driven flare models, by which one means models which require the magnetic energy dissipated during the reconnection process to be supplied continuously and from an external source throughout the flare. This type of model implies that the power from the external source supply be comparable to that necessary to keep the reconnection process going throughout the flare; (ii) In situ magnetic-energy storage models. This type of model requires the flare magnetic energy supply to be stored as a current passing through the energy release site, and while a power supply is necessary to build up the stored energy, this power supply is insufficient to actually drive the flare process on its own. The stored energy is sufficient to meet flare requirements. The purpose of this section is to discuss the second type of reconnection model and perhaps one of its fundamental difficulties.

The principal flare model that utilizes stored magnetic energy in the form of sheared magnetic fields and also releases this stored energy via reconnection is that proposed by Spicer (1977, 1981). This model, while heuristic in nature, introduced into the flare community many of the reconnection effects known to, or believed to, occur in laboratory devices, principally tokamaks and reversed field pinches. Various authors have developed "alternative models" to that proposed by Spicer but in reality these models are essentially the same, differing only in that different current profiles, or equivalently equilibriums, were utilized. As a result, different conclusions were derived concerning which reconnection modes would be important during a flare (e.g., Giachetti et al., 1977; Hood and Priest, 1979, 1981; Van Hoven, 1981). However, because all of these conclusions, as well as those of Spicer's, were in fact based on "assumed" equilibriums with little physical justification given for these assumptions, few firm conclusions can be drawn from this research concerning the applicability of this type of flare model.

Probably the most useful aspect of this type of flare model was that many new concepts concerning reconnection were introduced to the solar flare community, e.g., tearing modes in curvilinear geometries, non-constant " $\psi$ " tearing modes, tearing modes that are driven by an ideal MHD mode such as the  $m=1$  tearing mode (the so-called internal resistive kink), multiple tearing modes, the coalescence of magnetic islands, and magnetic braiding (stochastic fields; see Section

XII) that results from the mode coupling of different tearing modes (see Spicer, 1977, 1982). All of these mechanisms are known or believed to greatly enhance the rate of magnetic field dissipation via the reconnection process.

However, there may be one fundamental difficulty associated with all flare models of the type that use preflare magnetic energy storage. This difficulty is related to an observational constraint associated with the number of nonthermal electrons that need to be accelerated during the flare process in order to explain the observed flare initiated hard X-ray bursts. The difficulty arises when we recognize that those flare mechanisms that require pre-flare in situ storage of free magnetic energy can accelerate per unit time no more electrons than the number that carries the net current  $I$  associated with the stored energy. This argument (Spicer, 1983) is based on Faraday's law. The fundamental problem is the large disparity between, on the one hand, the magnitude of the allowed current consistent with the requirement that a given magnetic structure can stably store the necessary energy and, on the other hand, the magnitude of the current needed to explain the observed number of electrons. However, a suggestion by Haerendel (private communication, 1983) may overcome this difficulty by highly filamenting the current.

We must emphasize that the above argument, even if it proves correct, does not imply that the driven type of flare model cannot produce the electron flux necessary for the "nonthermal hypothesis." If the beam of electrons necessary under this hypothesis exists for a time shorter than the inductive decay time of the associated return current, very little new magnetic flux will be produced. Thus the self-magnetic field of the beam will remain small and the problems associated with the large self-fields of the bare beam are prevented.

We must also emphasize that in principle it may be possible to circumvent the difficulty we have just raised concerning the maximum electron flux,  $I/|e|$ , associated with inductive flare models. This can be accomplished by transferring the energy in this flux to a larger number of electrons with lower energy, e.g., by a beam-plasma instability. However, it must be recognized that such transfer of energy should provide a sizeable radiation signature detectable by present detectors.

In addition, it must be emphasized that an inductive flare model can indirectly accelerate a sufficient number of electrons per second to be consistent with the "nonthermal hypothesis" simply by driving a shock that in turn can give rise to a beam of  $10^{36}$ - $10^{38}$  electrons per second and thus a return current consistent with Faraday's law. In other words, while inductive flare mechanisms may not directly accelerate the electron numbers required in the manner originally proposed, they could accelerate the requisite numbers indirectly.

In summary, although in situ storage flare models satisfactorily explain the total energy release in a flare, they may fail to provide a direct explanation of that portion of the energy released as hard X-ray bursts. Model modifications of the type discussed above may help rectify that failure.

## X. MHD SIMULATION AND SMALL SCALE STRUCTURE

(W. H. Matthaeus)

### TURBULENCE IN RECONNECTION STUDIES

Much of our detailed knowledge of reconnection phenomena is based on computer simulation of the MHD equations. The task of accurately simulating solutions to fluid-like equations such as MHD in the nonlinear regime is not an easy one, mainly due to the appearance of turbulence in the solutions. MHD turbulence in two dimensions has been extensively studied (Kraichnan and



Montgomery, 1980) and is usually discussed in terms of Fourier wavenumber decomposition of the magnetic and velocity fields. Perhaps the single most important feature of MHD turbulence is the appearance of dynamically important excitations over a wide range of spatial scales. While the standard reconnection geometries do not lend themselves to a description in terms of homogeneous turbulence, for which the Fourier description is most relevant (Batchelor, 1982), a conservative approach to studying reconnection would be to allow for the possibility that the same sort of complex behavior may develop in the solutions.

The typical development of MHD turbulence, viewed as an initial-value problem, is as follows. Initially no fine scales are present. All of the energy is in large scale magnetic and/or velocity field Fourier modes. As time advances, magnetic and mechanical excitations spread to a wide range of spatial scales. The characteristic time scales for evolution of the system are the eddy-turnover times  $T_B = L/V_A$  and  $T_V = L/V$  where  $L$  is the scale of the energy containing fluctuations,  $V$  is the rms velocity, and  $V_A$  the rms Alfvén speed (Fyfe et al., 1977). The excited scales can usually be roughly divided into three regimes: the energy containing large scales, the smallest excited scales which constitute a dissipation range and an intermediate range in which energy spectra exhibit a near powerlaw dependence on wavenumber.

The system is also characterized by dimensionless mechanical and magnetic Reynolds numbers,  $R$  and  $R_m$  which are inversely proportional to the viscosity and resistivity, respectively. At large values of  $R$  and  $R_m$ , most of the dissipation occurs at scales very much smaller than the energy containing range. Accordingly, estimates of dissipation wavenumbers (Batchelor, 1982; Fyfe et al., 1977) increase with decreasing viscosity and resistivity. The spatial resolution required for accurate simulation is in no way related to the energy containing scale. It has been conjectured (Matthaeus and Montgomery, 1980) and to some extent numerically verified, that in the limit of very large  $R$  and  $R_m$  energy is dissipated at a nonzero rate independent of  $R$  and  $R_m$ . This is similar to the conjecture of Kolmogorov (1941) and Oubukoff (1941) which is widely accepted in the hydrodynamic turbulence community (e.g., Batchelor, 1982, Chapter 6).

Turbulence at high Reynolds numbers involves a large number of degrees of freedom (Fourier modes) and this number increases with increasing Reynolds number. Accurate simulations must retain an adequate number of Fourier modes (or their equivalents) and must freely allow among them the transfers of excitation implied by the MHD equations. There are a number of ways that simulation design can run afoul of the guidelines.

"Numerical diffusion" (Potter, 1977) afflicts many finite difference schemes. Use of such schemes may dissipate energy (or some other quantity) by both numerical and physical processes, making it difficult to determine how processes such as reconnection depend on the physical dissipation coefficients.

Some techniques (e.g., Sato and Hayashi, 1979) make use of neighboring cell filtering methods to smooth numerically calculated time derivatives, ostensibly to stabilize the method. This is a questionable approach when simulating reconnection which delicately depends on the formation of small scale structures such as a diffusion region, since filtering suppresses the transfer of excitation to high wavenumber in ways that are likely to be inconsistent with MHD. Large-scale dynamical features may be qualitatively reproduced, but information about rates and dependence on Reynolds numbers is probably lost.

The use of artificial dissipation terms explicitly added to numerical MHD methods is also not uncommon. These are usually either of the "superdissipation" type (Meneguzzi et al., 1981) wherein the extra dissipation terms are effective only at very high wavenumber, or of the "anomalous" dissipation type (Sato and Hayashi, 1979) which may increase dissipation when the electric current exceeds specified thresholds. The justifications for these approaches are quite

different and both must be interpreted with care. Superdissipation (Basdevant, 1981) techniques attempt to model the unretained very high  $k$  modes as a sink of energy and are of limited utility since a fairly large unmodified intermediate wavenumber range must be present to preserve accuracy. Anomalous dissipation is motivated by considerations outside the scope of MHD.

Another important issue in MHD simulations is that of numerical reconnection and the preservation of ideal MHD topological (pointwise) invariants in codes when resistivities are set to zero (Matthaeus and Montgomery, 1980). It is computationally impossible to preserve pointwise conservation laws for arbitrary initial conditions and arbitrarily long times. Codes treating nonlinearities in an internally consistent manner avoid numerical reconnection in cases of physical interest by restricting the values of Reynolds numbers used. When resistivity and viscosity are set to zero, accurately modeled nonlinearities should cause spectral transfer to the limit of resolution in a few turnover times, leading inevitably (and appropriately) to numerical reconnection. This phenomenon is in no way connected with dissipation of energy. Concern with maintaining pointwise conservation laws has sometimes led simulators to introduce some of the numerical artifices discussed above to delay or prevent numerical reconnection by suppressing transfer to small scales.

Some numerical models include much finer spatial resolution in the direction across the initial current sheet than in the direction along it. There is a possibility that a subtle problem is inherent in these anisotropic representations for those MHD solutions in which turbulence is expected. In incompressible MHD, excitations at wavenumber  $\underline{k}$  are directly coupled to those at wavenumbers  $\underline{r}$  and  $\underline{p}$  only when  $\underline{r} + \underline{p} = \underline{k}$ . The coefficients of the nonlinear couplings (Fyfe et al., 1977) vanish when  $\underline{p}$  and  $\underline{r}$  are parallel. On average the most efficient coupling occurs when  $\underline{r}$  and  $\underline{p}$  differ in magnitude by a ratio of about  $\sqrt{2}$  and are about  $45^\circ$  apart in angle. To achieve rapid reconnection, the current sheet must become much thinner than its initial width. High  $k_y$  modes must be excited by nonlinear interactions, where the  $y$  direction is across the current sheet. If very much higher resolution is retained in  $y$  than in  $x$  (the direction perpendicular to  $y$ ), a substantial number of the wavenumber triads which pump the high  $k_y$  modes are excluded. To efficiently excite high  $k_y$  modes,  $k_x$  modes must be retained which do not differ from  $k_y$  in order of magnitude. At this writing, we do not know of any systematic study which evaluates this effect. Highly anisotropic representations may underestimate turbulent thinning of the current sheet near X-points.

The type of computational restrictions discussed here are expected to be in effect unless special initial conditions, perturbations, or forcing mechanisms act to prevent turbulence from appearing.

## RECONNECTION SIMULATIONS USING THE SPECTRAL METHOD

Spectral-method simulation of the evolution of a periodic sheet pinch (Matthaeus and Montgomery, 1981; Matthaeus, 1982) is a somewhat nonstandard approach to studying the reconnection process. The two-dimensional MHD equations are solved numerically in a periodic box initially containing two sheet currents (to achieve periodicity) using the Orszag-Patterson spectral method (Orszag, 1971). The use of this simulation geometry prohibits identification with any particular physical system, but allows many of the features of reconnection to be studied. The principal advantages of the numerical scheme are the absence of aliasing and numerical dissipation errors and direct control of spatial resolution.

These simulations have shown the development of a reconnection zone near magnetic X-points which form along the initial current sheets. The dynamics are triggered by broad band low level noise rather than by symmetrized eigenmode perturbations. Plasma flows into the reconnection

zone from the strong field sides. Outflow along the current sheet is at a substantial fraction of the Alfvén speed. Magnetic field lines are continuously reconnected, eventually forming large magnetic islands. Enhanced dissipation occurs at X-points due to the formation of electric current filaments within the reconnection zone. A quadrupole-like distribution of vorticity is self-consistently generated around current filaments, augmenting the characteristic flow pattern. The vorticity is generated by the curl of the Lorentz force which becomes large near the X-point corners. This effect is absent in Parker-Sweet reconnection and is probably the nonsteady self-consistent analogue of the vorticity sheets (shocks) in the Petschek model (see Figure 1-20). The appearance of intense current filamentation, quadrupole-like vorticity distributions, and bursts of reconnection activity in these simulations indicate that MHD reconnection zones are highly structured both spatially and temporally. It is likely that such structure enhances reconnection to rates above those of Parker and Sweet by narrowing the reconnection zone in a way that is reminiscent of the Petschek model, but without the need for an externally supported flow field.

A simplified scaling discussion has been given (Matthaeus, 1982) to demonstrate that fine structures in the reconnection zone might plausibly result in rates of reconnection which do not vanish for large magnetic Reynolds numbers. For this to occur, the volume of the reconnection zone must shrink as  $R_m$  increases, while the intensity of the current increases so that the ohmic dissipation rate remains constant. This is the same limiting behavior required for "selective decay" to occur in MHD (Matthaeus and Montgomery, 1980). It is analogous to the Kolmogorov-Oubukoff conjecture concerning hydrodynamic turbulence.

Selective decay occurs through coalescence of magnetic islands. If a dissipative MHD simulation begins with 10 or so randomly positioned islands, turbulence leads ultimately to a state with just two large islands. At every stage of the turbulent decay leading to this final state, reconnection zones form between adjacent magnetic islands, exhibiting the same type of current and vorticity structures found in sheet pinch simulations. The enhanced ohmic heating produced by this continual reconnection in a turbulent MHD fluid has been proposed as a possible source of heating of the solar corona. There is an intriguing possibility that global turbulent MHD turbulence processes and reconnection rates are integrally related. Recently, Frisch et al. (1983) have investigated the appearance of small-scale structures near neutral points in two-dimensional MHD-turbulent flows. They use very high resolution codes ( $256^2$  and  $512^2$ ) with zero dissipation and run for short times. They define a "smallest excited scale" which monotonically decreases in time, indicating that dissipation alone can set a lower limit to the scale size of structures appearing near X-points. Recently, we have simulated the MHD sheet pinch at several values of  $R$  and  $R_m$  up to 1000. The most important result is that the time required for growth of "reconnected" magnetic islands does not increase as  $R_m$  is increased. The reconnection zones are observed to decrease in size as  $R_m$  is increased but it has not yet been conclusively shown that the scaling suggested above obtains.

## CONCLUSION

The overall importance of reconnection depends on how rapidly it occurs compared to other plasma dynamical processes. Historically, reconnection time scales have been thought to be considerably slower than eddy turnover times or Alfvén transit times. In the Parker-Sweet model and in tearing-mode theory, the characteristic reconnection times tend to zero algebraically as the magnetic Reynolds number goes to infinity. In astrophysical applications, the magnetic Reynolds numbers are so large that reconnection according to this scaling becomes negligibly slow in comparison with other dynamical processes. Fine-scale structures appearing near X-points may increase rates of reconnection above the usual estimates, leading to very rapid "turbulent reconnection."

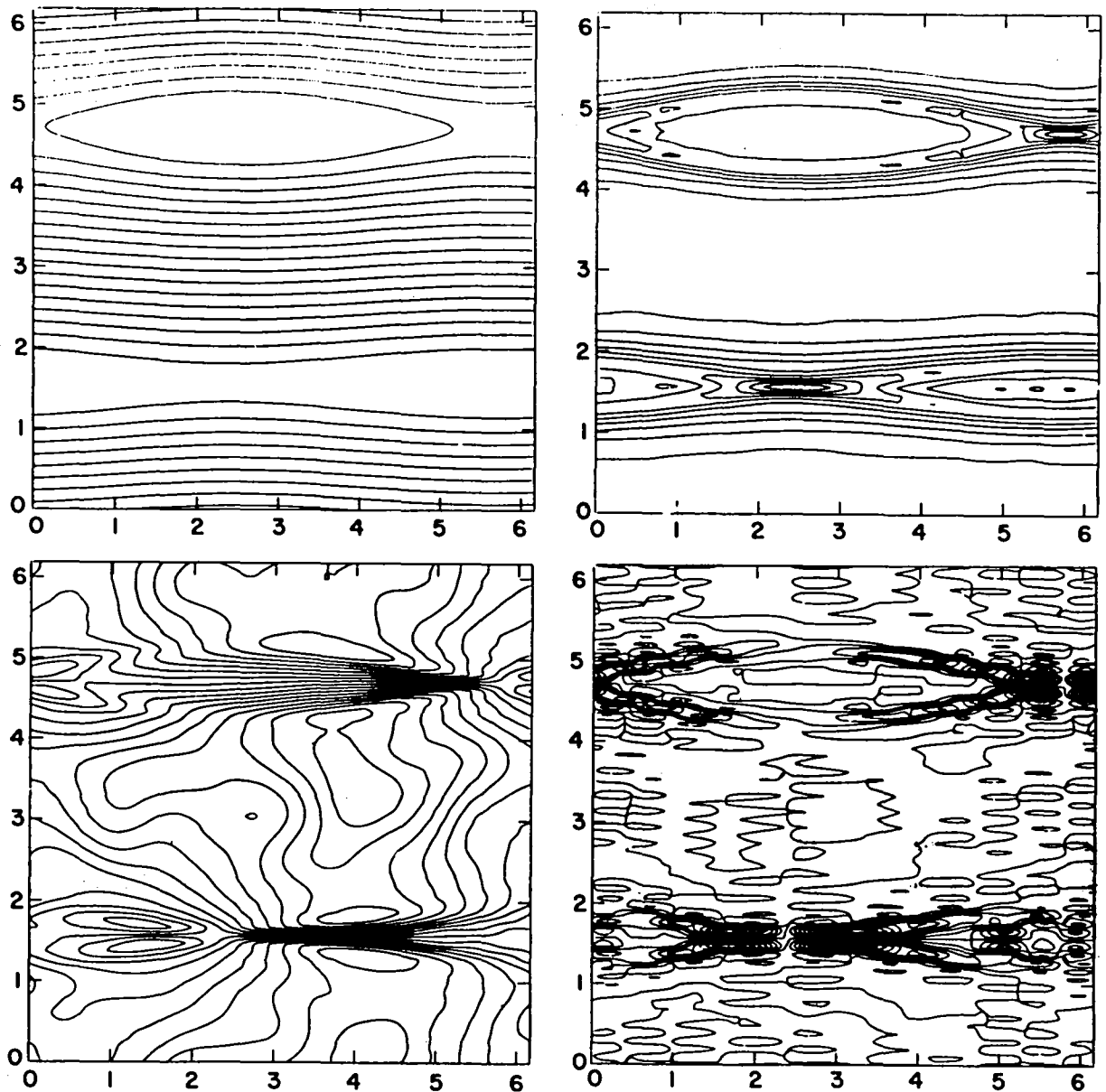


Figure 1-20. Magnetic field lines (upper left), fluid streamlines (lower left), contours of constant electric current density (upper right) and contours of constant vorticity (lower right), all after 21.4 Alfvén transit times of evolution. The field lines indicate that magnetic islands have grown due to the dynamical phenomena occurring at X-type neutral points. Current filamentation causes enhanced ohmic dissipation near X-points. The speed of plasma jetting from the X-point region reaches about 20% of the Alfvén speed in this simulation. The flow pattern is associated with quadrupole-like distributions of vorticity which spontaneously appear near the X-point region. The simulation is unforced. Mechanical and magnetic Reynolds numbers are both 400.

# XI. LABORATORY OBSERVATIONS OF NEUTRAL SHEETS

(R. L. Stenzel)

The UCLA reconnection experiment to be discussed here (Stenzel et al., 1983a, Gekelman et al., 1982) is shown schematically in Figure 1-21. It employs a linear discharge plasma column produced with a large diameter (1 meter) cathode. Detailed plasma diagnostic tools are employed in conjunction with a fast digital data acquisition system. Time and space resolved probe measurements of the magnetic fields, electric fields, macroscopic plasma parameters ( $n, T, v_d$ ) and microscopic properties [ $f(\vec{v})$ ] are performed. The digital data processing is performed with 100 MHz, 32k digital-to-analog converters, on-line computer with array processor and a data link to a Cray computer.

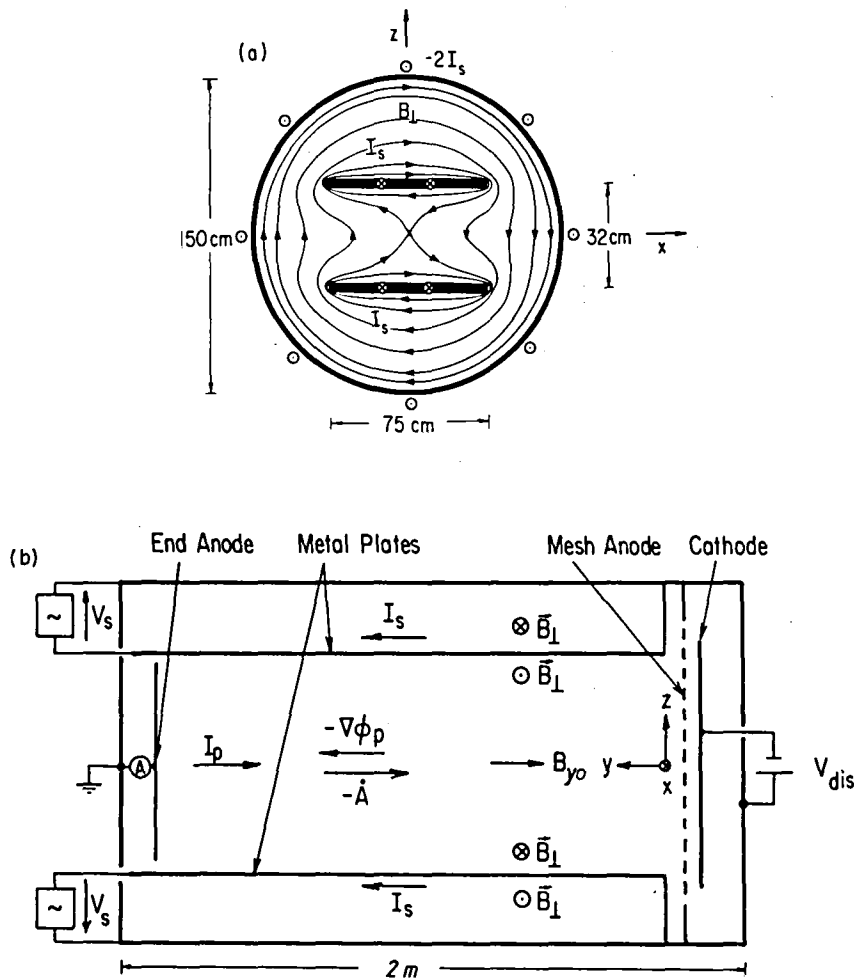


Figure 1-21. Schematic picture of the experimental arrangement. (a) Cross-sectional view showing parallel plate electrodes with pulsed currents  $I_s$  and magnetic field lines  $\vec{B}_{\perp}$  without plasma. (b) Side view of the device with main electrodes, currents ( $I_p, I_s$ ) electric fields ( $\vec{E} = -\vec{A} - \nabla\phi_p$ ), and magnetic fields ( $\vec{B} = \vec{B}_{\perp} + \vec{B}_{y0}$ ). The coordinate system common in magnetospheric physics has been adopted where  $y$  is along the neutral line (device axis),  $x$  is along the horizontal neutral sheet, and  $z$  is normal to the sheet.

Typical electron densities and temperatures in the experiment are  $n_e \approx 10^{12} \text{ cm}^{-3}$  and  $kT_e \approx 10 \text{ eV}$  with an initially axial magnetic field  $B_y \approx 10 \text{ G}$ . The plasma is uniform, quiescent, nearly collisionless, and highly reproducible in pulses of duration  $t_p \approx 5 \text{ msec}$ , repeated every  $t_r \approx 2 \text{ sec}$ . After establishing the plasma, a time varying transverse magnetic field ( $B_z \approx 10 \text{ G}$ ) is applied. Its topology in vacuum contains an X-type neutral point along the axis of the device. In the presence of the plasma, electron currents ( $I_p \approx 1000 \text{ A}$ ) are induced which flow preferentially in regions of  $B_z = 0$  but are sufficiently large to modify the applied field topology to that of a neutral sheet during the period of rising magnetic fields ( $0 < t < 100 \mu\text{sec}$ ). The current sheet thickness scales with the collisionless electron skin depth  $c/\omega_{pe}$  ( $\Delta z \leq 10 c/\omega_{pe}$ ) and is much smaller than the ion Larmor radius ( $r_{Li} \approx 30 \text{ cm}$ ). The long, thin ( $\Delta x \approx 50 \text{ cm}, \Delta z \approx 5 \text{ cm}$ ) sheet normally exhibits no tearing instabilities (Furth et al., 1963) or slow shock structures (Petschek, 1964). The magnetic Reynolds number based on the classical collisional resistivity is  $R_m \approx 10-100$  depending on the ion mass and B-fields chosen. For uniform plasmas and modest current densities ( $v_{de}/v_{te} \leq 0.1$ ) a macroscopically stable current sheet can be maintained in which reconnection takes place essentially as described in the Section I and shown in Figure 1-1.

## CURRENT DISRUPTION

In the experiment, the neutral sheet current terminates on a large grounded end plate. In order to raise the current density in the middle of the neutral sheet, the central portion (6 cm x 13 cm) of the end plate (32 cm x 75 cm) is isolated and biased positively ( $0 < V_a \leq 30 \text{ V}$ ). When monitoring the collected current to this center electrode ( $I_a$ ), we find that at increasing current levels ( $v_{de}/v_{te} > 0.3$ ) spontaneous disruptions occur (Stenzel et al., 1983a) which can result in a complete current switch off. At low current levels the waveform is given by the smooth envelope over the current peaks. By simultaneously monitoring the current to the surrounding large end plate we conclude that the disruption does not lead to a current loss but to a redistribution from the center of the neutral sheet to the sides. The cause for this current break-up has been inferred from detailed diagnostics of the local plasma properties.

During the disruption, the plasma potential rises in the perturbed current channel to a value much larger than the dc potential applied to the plate. Simultaneously, the particle density decreases. These phenomena have a finite extent along the X line. In particular, an abrupt drop of the plasma potential by  $\approx 30 \text{ V}$  is observed in an axial distance  $\Delta y \approx 5 \text{ mm} \approx 100 \lambda_D$ . Thus, in the location where the current is disrupted, a potential double layer is formed. The large positive plasma potential can only be understood in terms of the circuit properties (Alfvén, 1977; Spicer, 1982; Smith, 1982). The extended current path has a distributed inductance  $L$  which, in the presence of current changes, is responsible for a loop integrated inductive voltage  $L \, dI/dt$ . This voltage drops off inside the plasma in the region where the current opens. Both the local plasma properties and the circuit inductance determine the resultant behavior. The positive plasma potential in the perturbed current channel causes an expulsion of ions and hence a density and current drop. However, the current decrease in the inductive circuit reinforces the electric fields inside the plasma. The interaction leads to a negative current-voltage characteristic which results in the current disruption.

In this experiment, the onset of anomalous resistivity is neither the cause for current disruptions nor the mechanism for energy conversion from fields to particles (Baum et al., 1978). At the double layer, particle beams are generated, i.e., magnetic energy  $LI^2/2$  is directly converted to kinetic energy of the particles. Subsequent beam-plasma instabilities result in microinstabilities which in turn transfer beam energy into waves and heat. For example, the electron beam excites Langmuir waves ( $\omega > \omega_{pe}$ ) which are observed as bursts of microwave emission at each current disruption. At the low potential side, energetic ions and low frequency turbulence are observed.

The present experimental geometry makes a direct comparison with the Earth's tail current system or a solar loop difficult. Nevertheless, there are many common qualitative features. The tail current disruption and redirection during substorms is one example. The formation of auroral potential structures and the present double layer formation may also have common features. The particle acceleration mechanism in a double layer is well known from auroral physics but it is worth considering that these structures arise from inductive voltages in large current systems in which the energy is stored well outside the region of dissipation. Finally, the observed plasma wave emissions during current disruptions may be compared to type III solar bursts. Based on the successful circuit analysis of the laboratory observations, one should attempt global circuit modeling in space plasmas as well.

## MICROINSTABILITIES

A rich spectrum of waves can be excited by anisotropies of the particle distributions in a current sheet. Current driven ion-acoustic instabilities have been identified in previous experiments (Overskei and Politzer, 1976; Stenzel and Geikelman, 1981). Here, we concentrate on electromagnetic modes above the lower hybrid frequency ( $\omega_{lh}^2 \approx \omega_{ce}\omega_{ci}$ ) observed in a plasma with comparable magnetic and particle energy densities and in the absence of current disruptions.

Magnetic fluctuations are recorded with two movable magnetic probes each containing orthogonal loops to detect the vector components  $\vec{B}_\perp = (B_x, B_y)$ . Since the reconnection process is time varying, i.e., nonstationary, a large ensemble of traces  $B_i(t)$ ,  $i=1\dots 80$ , is recorded digitally and statistical properties derived by ensemble averaging. Figure 1-22a shows the typical frequency spectrum of magnetic fluctuations in the range above the lower hybrid frequency ( $f_{lh} \approx 100$  KHz). Peak amplitudes are of order  $\delta B_\perp / B_{y0} \approx 10\%$  and the spectral coefficient is approximately  $\alpha \approx 2-3$  in the formula  $\delta B_\perp \propto \omega^{-\alpha}$ . Figure 1-22b shows an example of the cross power spectrum at  $\omega/2\pi = 1$  MHz in the plane  $(k_x, k_y)$ . Even at a single frequency, a broad range of modes propagating at various angles with respect to the dominant axial magnetic field  $B_{y0}$  is observed. The wave which satisfies these properties is the oblique whistler wave. In the frequency regime under consideration ( $\omega_{lh} < \omega < \omega_{ce}$ ), the phase velocity cone ( $\cos \theta \approx \omega / \omega_{ce} \approx 0.03$ ) permits wave propagation at essentially any angle  $|\theta| \leq 90^\circ$  with respect to the magnetic field but with different wavenumbers  $k(\theta) \approx (\omega/c)(\omega_{pe}^2 / \omega \omega_{ce} \cos \theta)^{1/2}$ . The observed wavenumbers are in the range expected for the experimental plasma parameters. Further confirmation of this mode identification has been obtained from polarization measurements.

The results given above show that in a carefully diagnosed and controlled laboratory plasma, microinstabilities can be examined in great detail. Whistler-mode wave noise is by no means the only microturbulence present in current sheets: ion sound modes and electron plasma modes are observed as well. One of the difficulties in the investigation of microinstabilities is to determine which of the competing instabilities dominates in modifying the plasma properties by wave-particle interactions. This requires a comprehensive study of all instabilities as well as simultaneous particle measurements.

## ELECTRON DISTRIBUTION FUNCTIONS

Macroscopic plasma properties are derived from the microscopic phase space distribution function  $f(\vec{v}, \vec{r}, t)$ . A direct measurement of this function is the ultimate goal in plasma diagnostics. This task requires both a suitable particle detector and a high speed digital data processing system for analyzing a function of seven variables.

In the UCLA experiment, the particle detector consists of a velocity analyzer (0...150 eV) with narrow directional sensitivity ( $\Delta\Omega/4\pi \approx 10^{-3}$ ) accomplished by particle selection through a micro-channel plate (Stenzel et al., 1982; 1983b). The small detector ( $\sim 3$  mm radius) can be moved

in real space and at each position rotated through two orthogonal spherical angles  $\theta$ , and  $\phi$ , so as to measure the particle flux in three-dimensional velocity space. The measurements are time resolved to within  $\Delta t \approx 3 \mu\text{sec}$ . The distribution functions as assembled from individual measurements at typically 200 different directions  $(\theta, \phi)$ . Each differential flux measurement is repeated  $n \geq 10$  times so as to form ensemble averages and standard deviations. The large data flow is first handled with a fast array processor performing on-line data reduction and then transmitted to an off-site Cray computer for analysis and display.

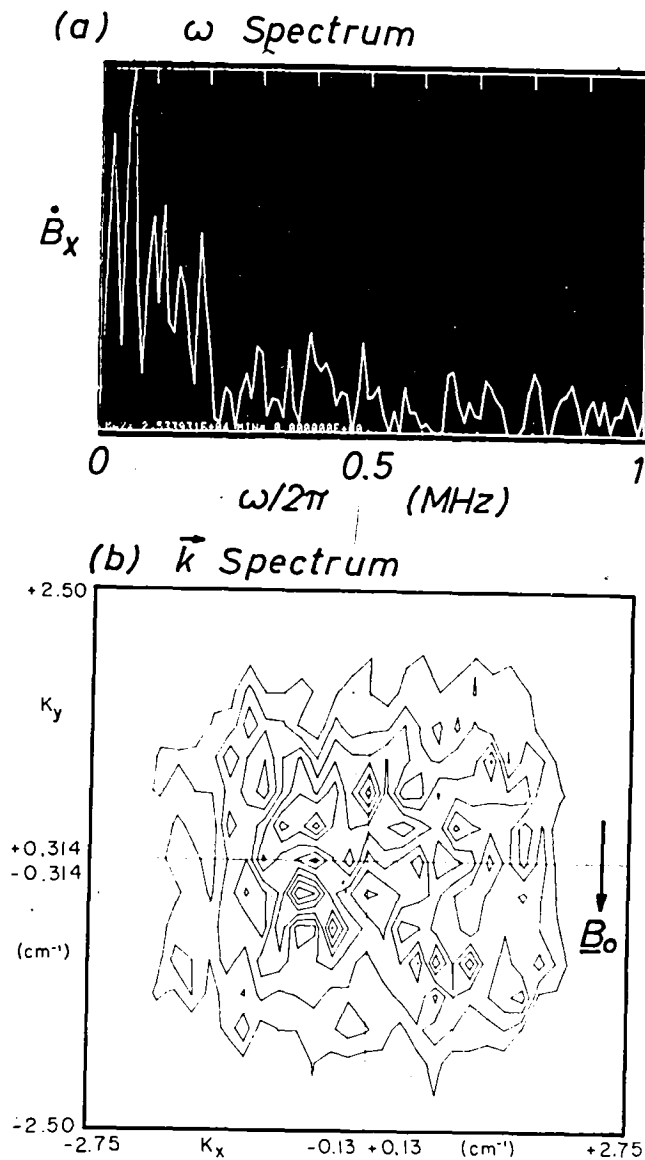


Figure 1-22. Frequency and wavenumber spectra of magnetic fluctuations in the neutral sheet. (a) Fourier analyzed probe signal  $\omega B_x(\omega)$ . Lower hybrid frequency  $\omega_{lh}/2\pi \approx 0.1$  MHz, electron cyclotron frequency  $\omega_{ce}/2\pi \approx 30$  MHz. (b) Two-dimensional spatial Fourier transform of the two-probe cross-spectral function

$$C_{12\omega}(\Delta \vec{r}, \Delta t) = \frac{1}{N} \sum_{i=1}^N B_{x1\omega}(\vec{r}, t) B_{z2\omega}(\vec{r} + \Delta \vec{r}, t + \Delta t) \text{ at } \omega/2\pi = 1 \text{ MHz.}$$



Figure 1-23 shows the measured distribution function  $f(v_x, v_y, v_z)$  displayed as surfaces of constant  $f$ . The maximum  $f$  value ( $f_{\max} \approx 10^{-14} \text{ cm}^{-6} \text{ sec}^3$ ) generates a surface near the origin; decreasing values produce surfaces of expanding radius in velocity space. Anisotropies cause deviations from a spherical shape. These are clearly visible on the surface  $f = 6 \times 10^{-16} \text{ cm}^{-6} \text{ sec}^3$ . The elongation along the magnetic field is produced by energetic electrons which run away in the inductive electric field along the separator. Figure 1-24 displays the electron distributions,  $\log f(v_{\parallel}, v_{\perp})$ , at two different positions: in the middle of the neutral sheet and at the edge. Runaway electrons are observed only inside the sheet; at the edge, the normal field component  $B_z$  prevents the free acceleration of electrons.

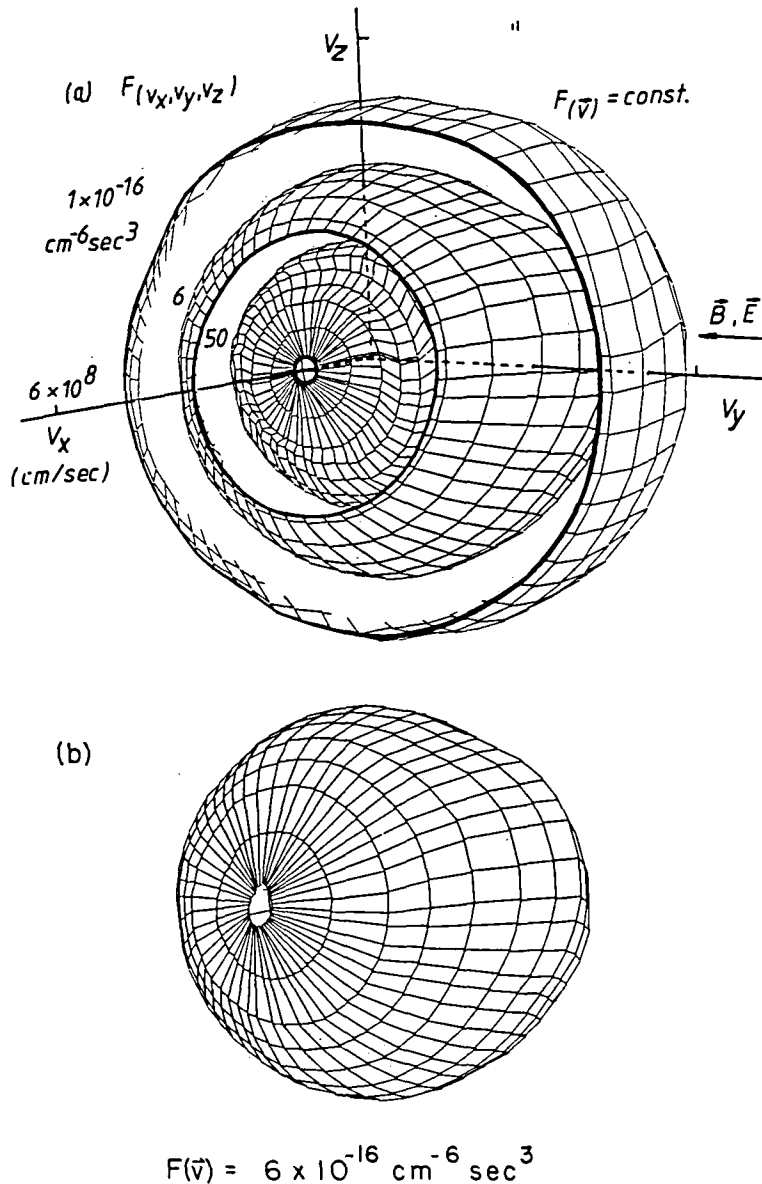


Figure 1-23. Display of the distribution function in three-dimensional velocity space as surfaces of constant  $f(v_x, v_y, v_z)$ . (a) Three nested surfaces cut for purpose of display. Radius expands with decreasing value of  $f(\vec{v})$ . (b) Complete display of the middle surface showing velocity space anisotropy (nonspherical shape) due to runaway electrons in a current sheet.

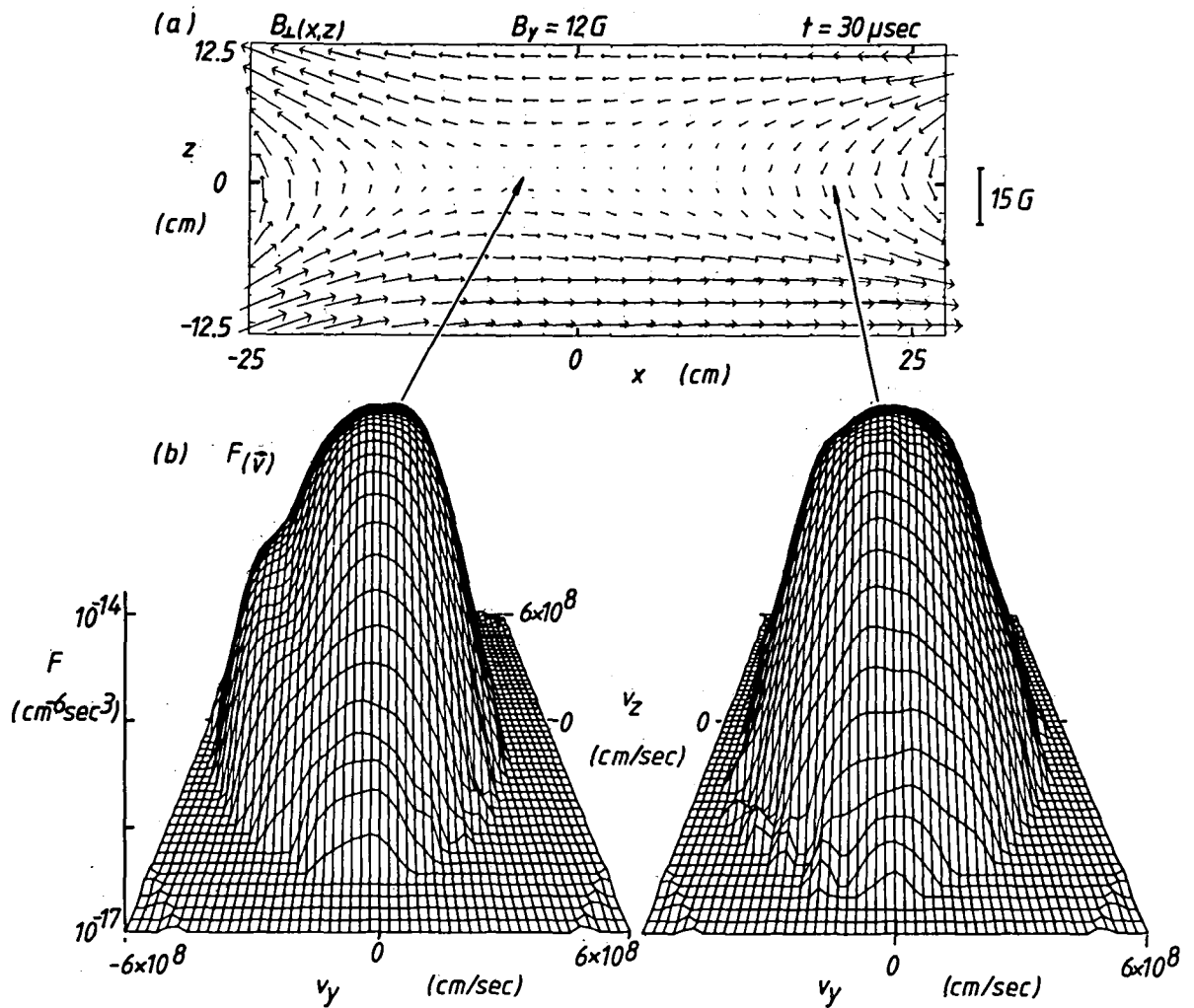


Figure 1-24. Spatial variation of the distribution function in the neutral sheet. (a) Magnetic field topology with locations where  $f(\vec{v})$  was measured. (b) Electron distribution  $f(v_y, v_z)$  on a logarithmic scale with reference plane three orders of magnitude below the maximum. Note the presence of an energetic electron tail in the middle of the neutral sheet and absence of runaway electrons at the sheet edge.

The presence of runaway electrons in the neutral sheet is significant to both transport processes and instabilities. A few percent of runaway electrons carry more current than the drifting background population so that estimates of a classical resistivity based on the bulk plasma temperature become meaningless. Furthermore, spatial gradients in the tail population will cause inhomogeneities in the effective resistivity. Similarly, the energy transport is dominated by runaway electrons along the X line. Background plasma heating depends on the equilibrium processes in an anisotropic electron distribution. For small classical collision rates, micro-instabilities are effective in transferring energy from the free energy source via waves to the absorbing main population. Runaway electrons are a source of various microinstabilities, in particular, high frequency ( $\omega \geq \omega_{pe}$ ) electron plasma waves. Observations show that their damping is associated with bulk electron heating and conversion into electromagnetic radiation (Whelan and Stenzel, 1981, 1983).

The bulk electron temperature has been measured with high time resolution ( $\Delta t = 1 \mu\text{sec}$ ) using a novel array Langmuir probe (Wild et al., 1983). Significant temperature fluctuations ( $\Delta T/T \approx 10\%$ ) are observed in the neutral sheet. Heating appears to occur in bursts of short duration (2-5  $\mu\text{sec}$ ) generated in narrow regions (1-3 cm diameter) propagating across the magnetic field at approximately the sound of speed. These results are obtained from two-probe cross-correlation measurements of temperature fluctuations, an example of which is shown in Figure 1-25.

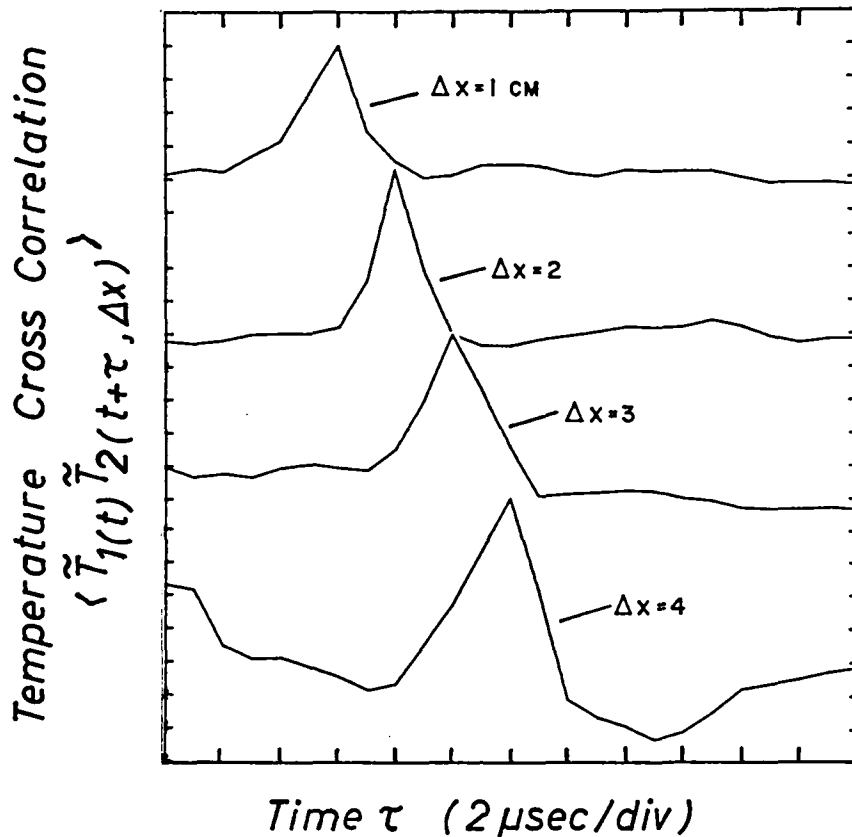


Figure 1-25. Two-probe cross correlation function of electron temperature fluctuations  $\tilde{T} = T - \langle T \rangle$  vs. delay time  $\tau$  at different probe separations  $\Delta x$ . Note the propagation of a heat pulse with cross-field velocity  $v_x = \Delta x/\tau \approx 5 \times 10^5$  cm/sec originating in the current sheet.

The knowledge of the microscopic heating and transport processes in collisionless turbulent plasmas is still very limited. During reconnection several low frequency and high frequency instabilities can be excited and interact with different parts of the anisotropic distribution. Careful wave and particle measurements are prerequisites for understanding the coupled processes. A controlled laboratory experiment is highly suited for this task. Based on such laboratory observations, one may suggest certain signatures of the diffusion region and then look for them in space where it has not yet been possible to locate the X-line region. These signatures include an enhanced level of magnetic turbulence coupled with energetic electrons moving along the separator.

Further attention should be paid to the role of the magnetic field component along the separator. It alters the particle dynamics, the relative contributions of electrons and ions to the current flow, the energy transport, and the nature of microinstabilities.

## XII. RECONNECTION IN TOKAMAKS

(W. Park)

In tokamak research (for an introduction, see Artsimovich, 1972), the magnetic-field topology plays an important role. Since the field is strong, the particles follow the field lines closely and the plasma confinement characteristics are largely determined by the field topology. Changes in this topology occur mainly as a result of plasma instabilities (tearing modes, etc.) (Furth et al., 1963; Rutherford, 1973; White et al., 1977).

In the present paper we examine the reconnection rates in two- and three-dimensional plasma evolution. In the former case, the two-dimensional symmetry ensures that all magnetic surfaces remain intact. For this situation, we have generalized the classical Sweet-Parker and Petschek reconnection-rate formulas to include the effects of the viscosity,  $\mu$ . In the latter case, regions of stochastic field line behavior can develop: the implications of this effect on tokamak reconnection are discussed.

### TWO-DIMENSIONAL RECONNECTION RATES

As the  $\beta$  value of tokamak plasmas increases, the kinematic viscosity  $\mu$  may become important in the reconnection process since, in rationalized emu units, the classical transport coefficients give  $\mu_1/\eta \approx (m_i/m_e)^{1/2}\beta$ . This viscous effect will also be important in space plasmas where  $\beta \sim 1$ . The Sweet-Parker and Petschek reconnection rates can be modified to include viscosity. The result is (Park et al., 1983):

$$\begin{aligned} M_A \psi_s &\sim \eta^{1/2} (1 + \mu/\eta)^{-1/4} && \text{(Modified Sweet-Parker)} \\ M_A \psi_s &\sim (1 + \mu/\eta)^{-1/2} && \text{(Modified Petschek)} \end{aligned}$$

where  $\dot{\psi}_s$  is the reconnected flux per unit time (which is equal to the reconnection electric field).

For ideally ( $\eta=0$ ) unstable tokamak plasmas, the reconnection rate from two-dimensional MHD simulations was found to satisfy the modified Sweet-Parker scaling. Figure 1-26 shows the evolution of magnetic surfaces in a simulation of a plasma unstable to the  $m=1$  ( $m$  signifies the  $e^{im\phi}$  component in a Fourier expansion) internal kink. This sequence of evolution is called a Kadomtsev-Monticello flip, or  $m=1$  flip (Kadomtsev, 1975). During the transition from the initial unstable equilibrium (a) to the final steady state (c), rapid reconnection (here defined by  $\dot{\psi}_s \sim \eta^k$  with  $k < 1$ ) takes place forming a magnetic island (b). A plasma state during this phase is shown

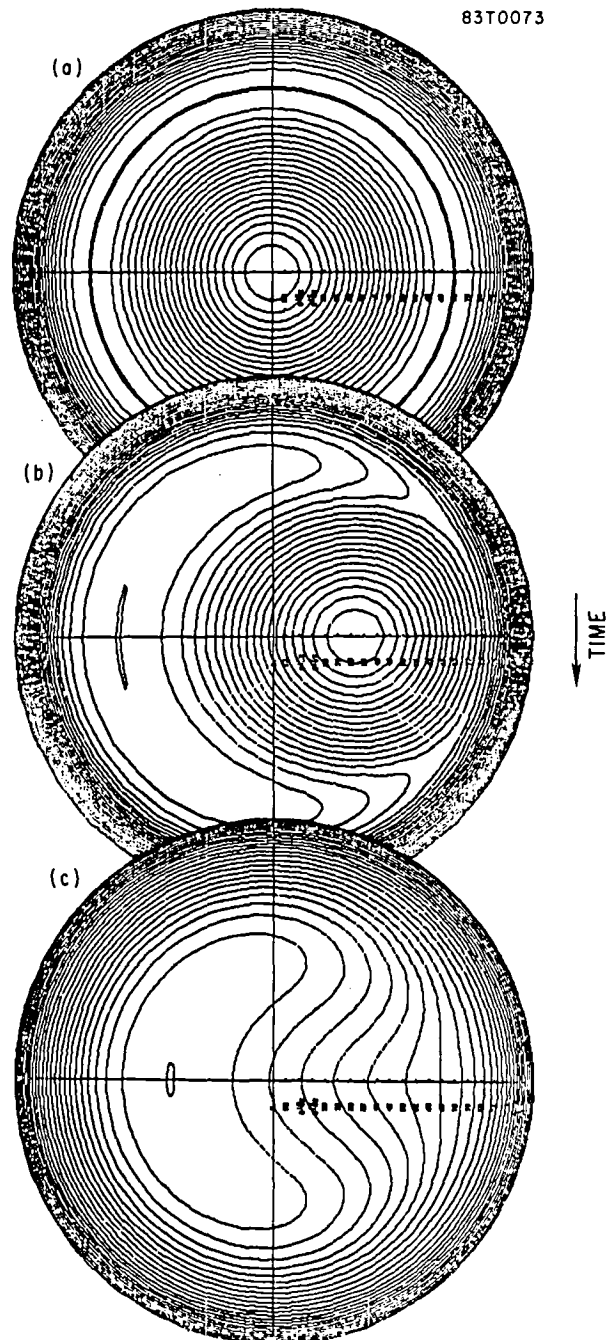


Figure 1-26. Time history of magnetic surfaces in an "m=1 flip" process. The magnetic field lines inside the singular surface where the X-point lies are reconnected to the field lines outside of that surface forming an m=1 island as shown in (b). In (c), all the field lines inside the singular surface are reconnected and the plasma has reached a steady state.

in Figure 1-27. Note the jump in  $B_\phi$  across the X-point in (b) and the long and narrow flow structure in (c) and (d). Figure 1-28 compares the numerical reconnection rate measured by the X-point current density,  $J_s$  ( $\psi_s = \eta J_s$ ); solid lines) to the modified Sweet-Parker scaling (dashed lines) in the regime where  $\mu \gg \eta$ . In the  $\mu \ll \eta$  regime, the current sheet shown in Figure 1-27 becomes unstable to higher  $m$  tearing modes and tears into multiple current sheets possibly leading to a turbulent region around the reconnection layer. A reconnection process driven by external coils has also been studied: the characteristics of the reconnection process were found to be essentially the same as in the ideally unstable case.

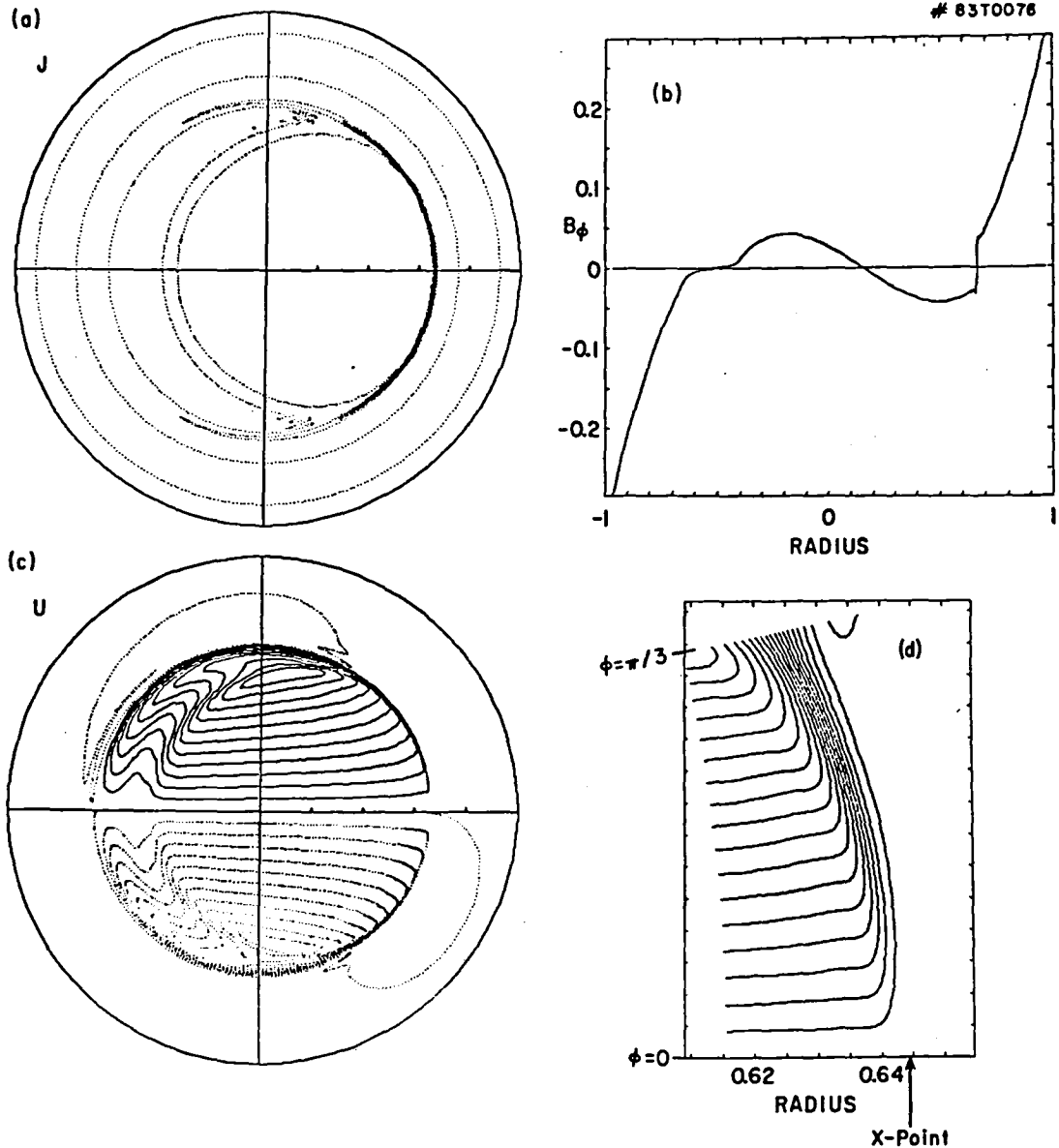


Figure 1-27. Profiles of plasma state during the reconnection phase (b) of Figure 1-26. A current sheet in the reconnection layer is apparent in (a), with the corresponding jump in  $B_\phi$  across the X-point shown in (b). Flow lines are shown in (c), with an enlarged (5 times more in  $r$  than in  $\phi$  direction) section around the reconnection layer shown in (d).

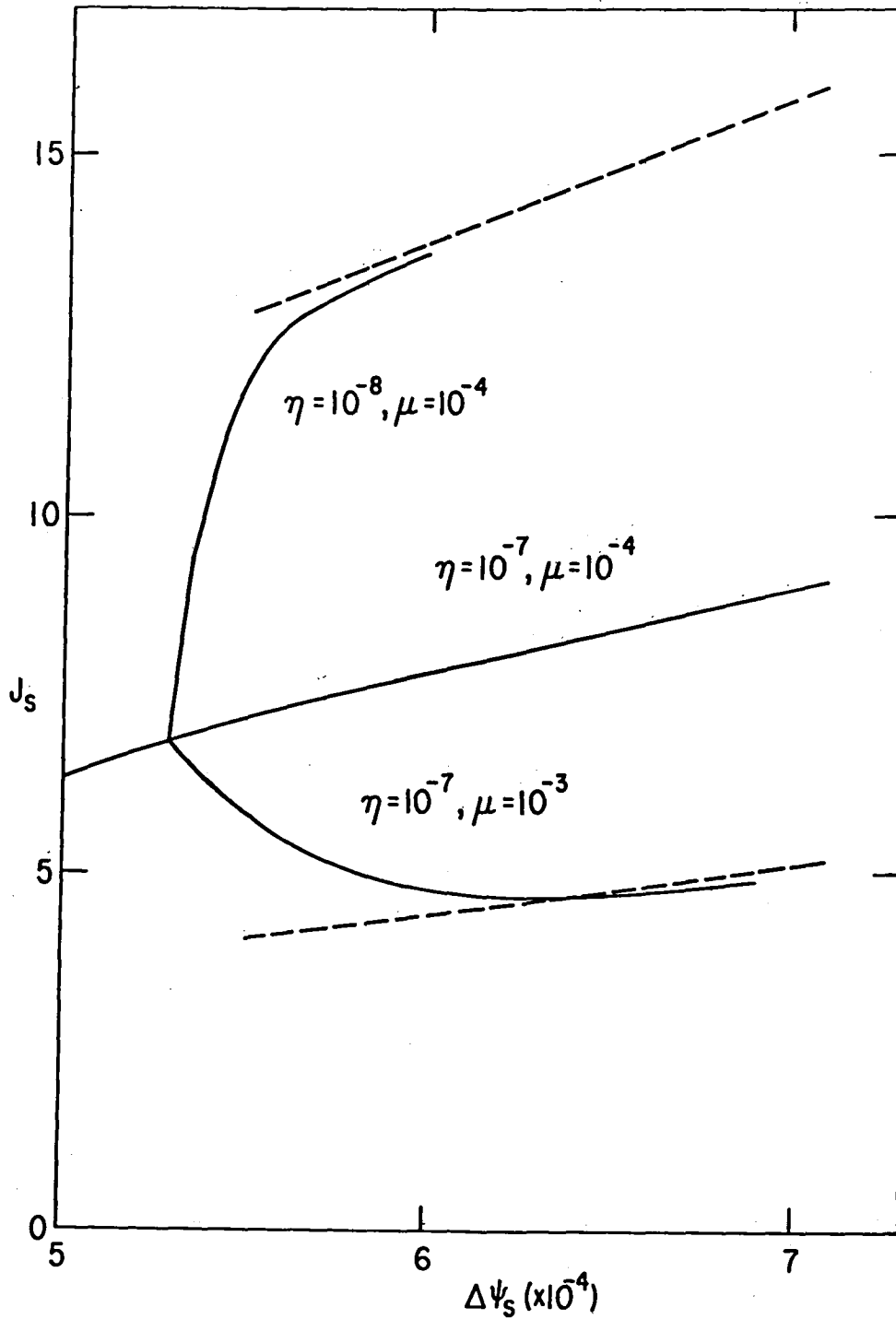


Figure 1-28. Current density  $J_s$  at the X-point is plotted against  $\Delta\psi_s$  which defines the plasma state. The results show good agreements between the simulation (solid lines) and the modified Sweet-Parker scaling (dashed lines obtained by applying the scaling to the middle solid line).

In the nonlinear evolution of purely resistive instabilities, i.e., instabilities which vanish for  $\eta=0$ , the reconnection rate was found to be slow (here defined as  $\dot{\psi}_S \sim \eta$ ). Figure 1-29 shows a state during the evolution of a plasma which is unstable to the  $m=2$  tearing mode. No jump in  $B_\phi$  is seen in (b). The current spike  $J_S$  in (c) was found to approach a finite value in the limit  $\eta \rightarrow 0$  which implies that the reconnection is indeed slow ( $\dot{\psi}_S \sim \eta$ ,  $J_S \sim \eta$ ).

The problems discussed above have closed boundaries. It would be of interest to study an open boundary problem where the reconnected flux can escape freely to see if it satisfies the modified Petschek scaling.

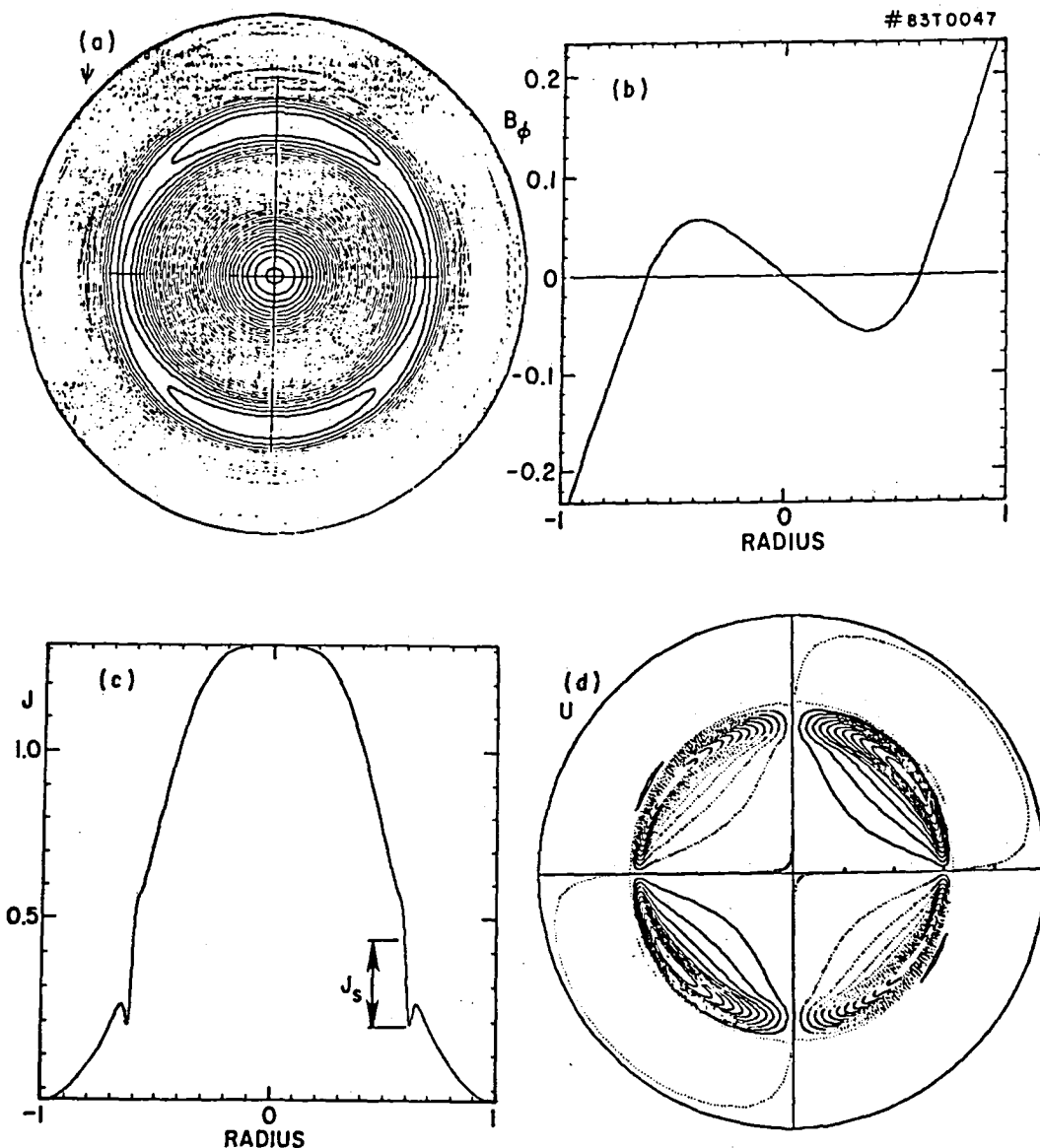


Figure 1-29. Profiles of a plasma state during a nonlinear evolution of the  $m=2$  "purely resistive" instability. No jump in  $B_\phi$  is visible in (b), and the  $J_S$  in (c) was found to reach a finite value as  $\eta \rightarrow 0$ , giving a slow reconnection (i.e.,  $\dot{\psi}_S \sim \eta$ ,  $J_S \sim \eta$ ).



### THREE-DIMENSIONAL SIMULATION OF TOKAMAK DISCHARGES

A three-dimensional initial-value MHD code is used at Princeton to simulate the nonlinear evolution of tokamak discharges (Park et al., 1981). During the simulation, the values corresponding to the diagnostic measurements are evaluated and compared to the experiment. Figure 1-30 shows an experimental soft X-ray emission measurement (center channel, BO) of a PDX tokamak discharge, while Figure 1-31a shows the details of signals (including the BO channel), during the first drop phase of the sawtooth curve in Figure 1-30 on an expanded time scale. As shown in Figure 1-31b, c, the corresponding signals from the simulation closely match the experimental data implying that the gross behavior of the actual plasma is correctly simulated. The extensive information available from the simulation has been used to explain many additional experimental phenomena.

In these simulations, the time scale was not matched with the experiment because a Lundquist number,  $S$ , i.e., a magnetic Reynolds number based on the Alfvén speed, smaller than the actual experimental value of  $S=10^7$  was used. With present-day computers, excessive time is required to perform a three-dimensional simulation of a complete plasma evolution using  $S=10^7$ . Thus, the time scales are deduced from the theoretical two-dimensional reconnection rates given above together with a simulation using a lower  $S$  value.

The drop phase of the soft X-ray signal in Figure 1-30 is believed to correspond to a  $m=1$  flip process (Waddell et al., 1978). From the modified Sweet-Parker scaling and a simulation with low  $S$  value, we can estimate the time scale of the  $m=1$  flip to be  $\sim 0.4$  msec which compares reasonably with the experimental value of  $\sim 1$  msec (from Figure 1-31a). A similar procedure using the Petschek scaling would give a time scale of  $\sim 1 \mu\text{sec}$  which is clearly too short. This application of the theoretical scaling assumes the actual three-dimensional reconnection rate to be close to the two-dimensional one. Also it is assumed that the effect of finite ion gyroradii on reconnection is small.

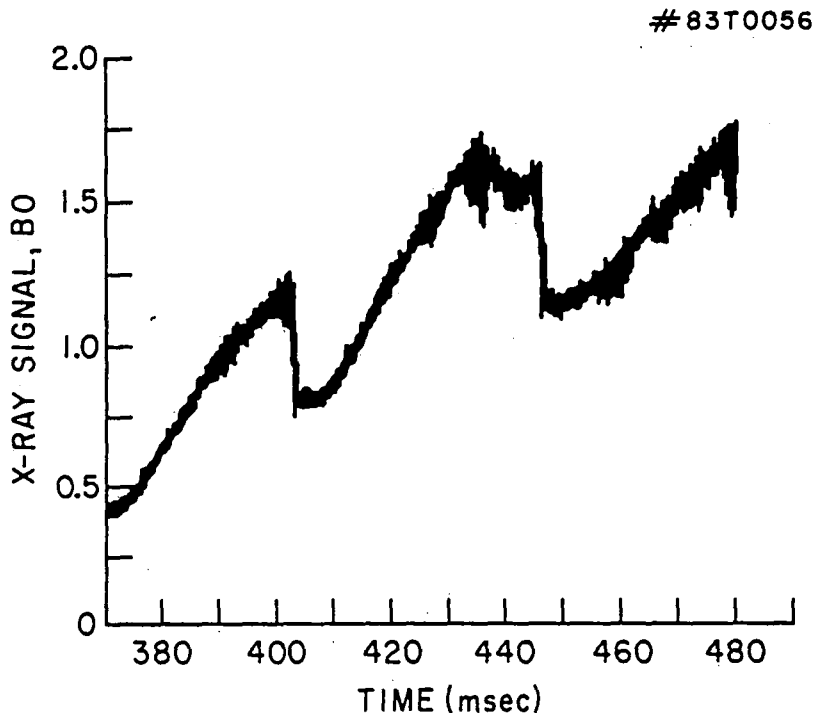


Figure 1-30. Experimental soft X-ray signals (BO channel) from a PDX tokamak discharge.

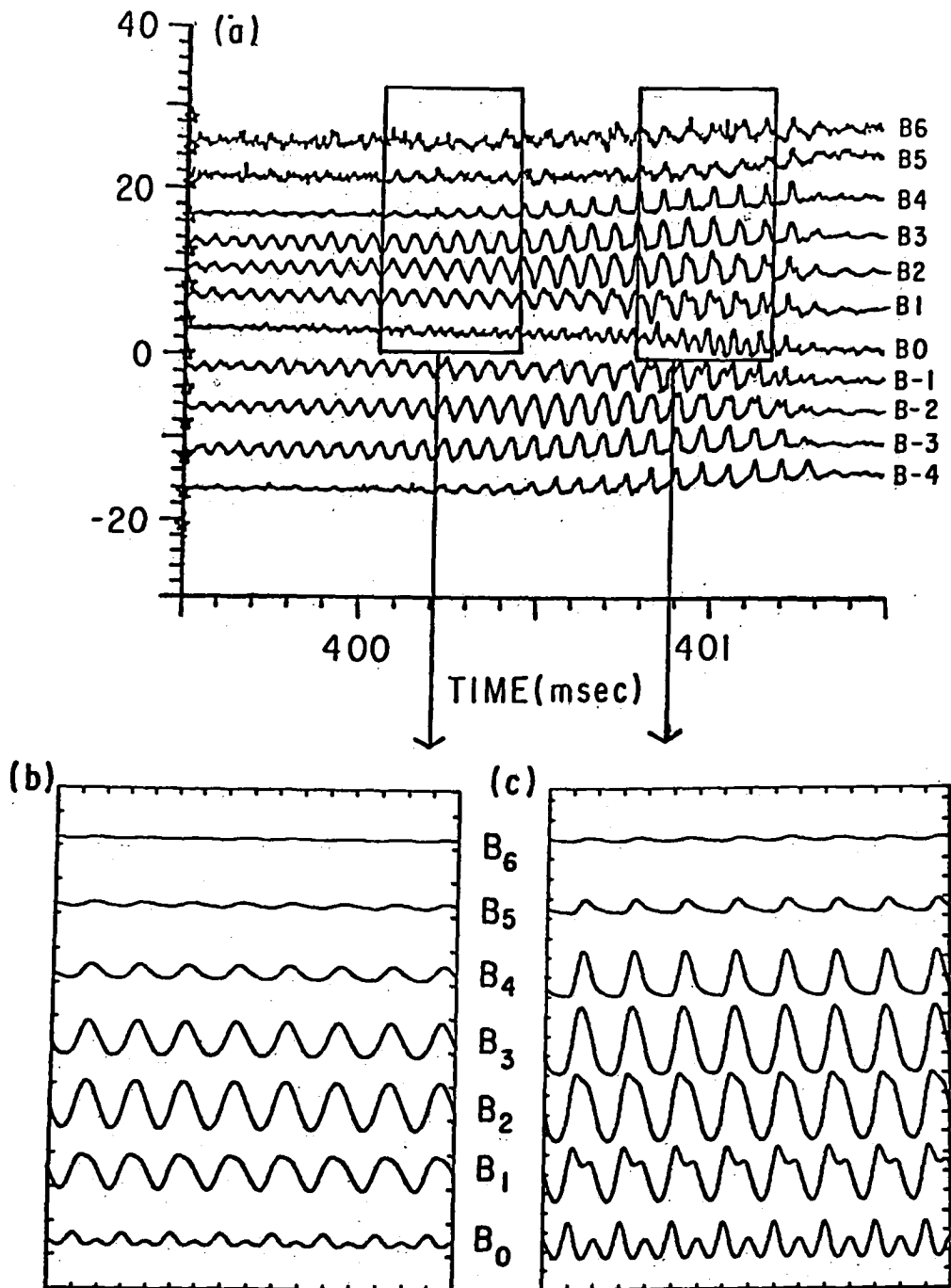


Figure 1-31. (a) Experimental soft X-ray signals from the first drop phase of the curve in Figure 1-30 on an expanded time scale (channel B0 as well as other channels). (b) and (c) are simulation results at  $\Delta B/B \sim 0.15\%$  and at  $\Delta B/B \sim 0.30\%$ , respectively.

### THREE-DIMENSIONAL RECONNECTION RATES

A field line entirely contained in a finite region (e.g., inside a torus) can either remain a line by closing onto itself, or it can trace a surface, or it can fill a finite volume. For any two-dimensional case, a flux function  $\psi$  exists such that a given field line lies in a  $\psi = \text{constant}$  surface. When the symmetry is broken, however, no flux function can be defined and a field line can fill a finite volume which is then called a stochastic field region. Figure 1-32 shows one such example. This figure is a field line map (Poincaré map) from the simulation described in the last section. A large  $m=1$  magnetic island is visible in the middle, surrounded by a stochastic region. In this three-dimensional problem, the helical flux function  $\psi$  used in the previous definition of the reconnection rate does not exist and the X-point is blurred into a region. A suitable three-dimensional definition of the reconnection rate is the maximum field-aligned electric field  $E_B = \eta \vec{J} \cdot \vec{B} / |B|$  in the X-point region.

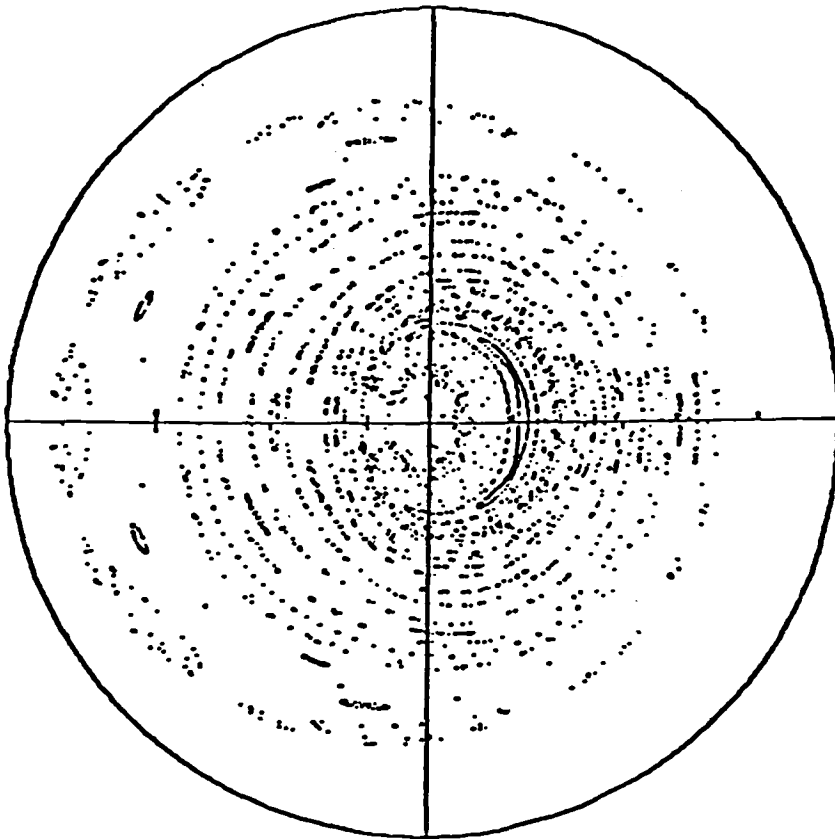


Figure 1-32. Poincaré map of magnetic field lines. Each point represents an intersection between a field line and the poloidal plane. The  $m=1$  magnetic island visible in the middle consists of a good surface region where a field line lies on a surface, and a stochastic region where a field line fills a finite volume.

In plasma confinement devices, the particle diffusion across a stochastic region or a magnetic island occurs effectively instantaneously. Thus, it is of much interest to evaluate the growth rate of a magnetic island or a stochastic region in various three-dimensional circumstances. This growth rate, of course, is governed by the reconnection rate. We have seen that two-dimensional reconnection in an ideally unstable plasma is rapid, satisfying the modified Sweet-Parker scaling, while reconnection in a resistively unstable but ideally stable plasma is slow ( $\psi \sim \eta$ ). Preliminary studies indicate that these results remain valid in three-dimensional reconnection.

Specific examples of the consequences of magnetic island growth in tokamak experiments are given below. The island growth in the  $m=1$  flip process causes the destruction of the original magnetic surfaces inside the  $m=1$  resonant surface, giving rise to a sudden temperature drop inside this surface as detected by the drop in the soft X-ray signal shown in Figure 1-30. The resulting temperature variation is called a sawtooth oscillation or a minor disruption (von Goeler et al., 1974). The growth of  $m=2$ ,  $m=3$ , ... islands destroys the magnetic surfaces in the outer region of the plasma where the resonant surfaces lie. The resulting stochasticity can explain the major disruption of tokamak discharges, which causes a sudden large loss of the plasma and sometimes terminates the discharge completely (Waddell et al., 1979). Schemes have been proposed to stabilize the dangerous  $m=2$  islands by inducing external current-driven islands near the  $m=2$  resonant surface. For this scheme to be feasible, the driven islands should be formed much faster than the growth of the  $m=2$  islands. This appears possible if the two-dimensional reconnection rates are valid. However, a more definite answer to this question must await further results concerning the three-dimensional generalization of reconnection.

## SUMMARY

In this section, we have presented the modifications to the Sweet-Parker and Petschek reconnection rates that are caused by the inclusion of viscous effects. The resulting rates have been compared to those obtained in two-dimensional computer simulations of tokamak reconnection with the following results. For resistively unstable but ideally ( $\eta=0$ ) stable plasmas, reconnection occurs at a rate ( $\psi \sim \eta$ ) slower than that predicted by the modified Sweet-Parker formula. For ideally ( $\eta=0$ ) unstable plasmas, reconnection occurs at the modified Sweet-Parker rate and this is also observed in tokamak experiments. The even more rapid rate given by the modified Petschek formula is not observed in the simulations. However, the Petschek rate may well be the appropriate one in many space applications where the boundary conditions are very different from those in a tokamak.

We have also discussed the results of three-dimensional computer simulations of reconnection in tokamaks. An important feature is that in a general three-dimensional plasma evolution, stochastic magnetic field regions can occur. In tokamaks such regions degrade the particle confinement. The implications of stochastic field regions in space plasmas, e.g., in the plasmoids formed in the geomagnetic tail, is an interesting and perhaps important topic for future study.

## XIII. CONCLUSION

(B. Sonnerup)

This study of magnetic field reconnection and its possible occurrence and importance in the magnetosphere and on the Sun has drawn, not only upon direct observations and measurements, but also upon available theory, upon computer and laboratory simulations, and upon examination of the process as it occurs in tokamaks. It has led to the following overall assessment.

The reconnection model has served, and continues to serve, as a successful organizing concept in magnetospheric as well as in solar physics. In the former area, recent spacecraft observations have helped move our capability to the point where direct quantitative comparisons between the predictions of simple reconnection theory and observations have become possible. These comparisons have provided impressive support for reconnection both in the geomagnetic tail and at the magnetopause but they have also demonstrated that the theory itself must be developed further if it is to serve as an effective guide in the future. Global as well as local computer models will undoubtedly be a principal means to achieve this goal, since it appears that global boundary conditions as well as the presence of local microscopic processes play an important role in the overall dynamics of magnetospheric reconnection. The success of computer simulations in helping to establish the occurrence and properties of reconnection in tokamaks serves as a useful illustration of this point.

The tokamak studies also show that reconnection can proceed even in circumstances where the magnetic field along the separator is the dominant field. Thus, any field-aligned current sheet or filament in the magnetosphere or elsewhere is a potential site for reconnection.

One of the most remarkable recent observational results is that magnetopause reconnection, which is thought of as a driven process, rarely occurs in a steady state. Rather it manifests itself in a patchy and highly time-dependent fashion. The theoretical aspects of these so-called flux transfer events are poorly understood and new theory as well as computer and laboratory simulation would be desirable. Flux transfer events promise to provide a major mechanism for the transfer of mass, momentum, and energy across the magnetopause leading, in an average sense, to a turbulent model of magnetopause reconnection. These observations also suggest that some kind of threshold, in addition to a southward magnetosheath field, may exist for the onset of magnetopause reconnection and that the magnetopause, either by accident or by design, operates in the immediate vicinity of the threshold. This description agrees in a qualitative way with laboratory terrella experiments which indicate the stable occurrence of closed as well as open front-side magnetopauses. These circumstances indicate that we have an outstanding opportunity to identify from spacecraft observations what are the critical onset and switch-off conditions for magnetopause reconnection. A similar opportunity exists for impulsive reconnection in the tail.

Recent three-dimensional time-dependent numerical simulations of magnetotail reconnection have confirmed the basic features of classical reconnection models. They have also duplicated a number of additional observed substorm-related signatures and they contain many new features which may form a basis for further detailed comparisons with observations. In particular, field-aligned currents and complex neutral-sheet structure form integral parts of the models.

We take particular notice of the incomplete state of observational and theoretical knowledge concerning the plasma processes near the separator line, i.e., in the diffusion region. Microinstabilities may play an important role in this region but, if so, it has not been established which instabilities are most important. Information concerning these processes may perhaps be obtained from laboratory experiments but it is not clear that such results can be translated directly to the magnetosphere application. The possibility also exists that in the tail and at the magnetopause resistivity provided by particle inertial and gyroeffects dominates. Since the processes in the diffusion region may be instrumental in determining the onset and switch-off criteria for reconnection, the development of a better understanding of these processes is of the essence.

In light of the absence of in situ measurements, and in light of the complicated nature of the solar magnetic field, it is not surprising that our knowledge of solar magnetic field reconnection is less developed than that of magnetospheric reconnection. Nevertheless, recent spacecraft

missions such as Skylab and SMM have provided a new, extremely valuable, data base for systematic studies of the reconnection process itself in its solar setting. Such studies should be vigorously pursued. On the theoretical side, computer models of flares, and models that draw upon our experience with reconnection in the magnetotail and in tokamaks offer promise. However, it must be remembered that the solar environment has many important features not present elsewhere. The possibility of coronal heating by small-scale turbulent reconnection is an outstanding example of a problem where recent numerical simulations may provide crucial insights.

#### XIV. REFERENCES

- Aggson, T. L., P. J. Gambardella, and N. C. Maynard, Electric field measurements at the magnetopause: 1. Observations of large convective velocities at rotational magnetopause discontinuities, J. Geophys. Res., **88**, 1983 (in press).
- Alfvén, H., Some properties of magnetospheric neutral surfaces, J. Geophys. Res., **73**, 4379, 1968.
- Alfvén, H., Electric currents in cosmic plasmas, Rev. Geophys. Space Phys., **15**, 271, 1977.
- Artsimovich, L. A., Tokamak devices, Nucl. Fusion, **12**, 215, 1972.
- Axford, W. I., Magnetospheric convection, Rev. Geophys., **7**, 421, 1969.
- Babcock, H. W., The topology of the sun's magnetic field and the 22 year cycle, Astrophys. J., **133**, 572, 1961.
- Basdevant, C., B. Legras, R. Sadourny, and M. Beland, A study of barotropic model flows: intermittency, waves, and predictability, J. Atmos. Sci., **38**, 2305, 1981.
- Baum, P. J., A. Bratenahl, and G. Kamin, Current interruption and impulsive flux transfer solar flare models, Astrophys., **226**, 286, 1978.
- Baum, P. J., and A. Bratenahl, The laboratory magnetosphere, Geophys. Res. Lett., **9**, 435, 1982a.
- Baum, P. J., and A. Bratenahl, Rapid magnetic flux transport in laboratory and space plasmas: the effect of tensor conductivity, EOS, Trans. Am. Geophys. Union, **63**, 408, 1982b.
- Batchelor, G. K., Theory of Homogenous Turbulence, Cambridge U. Press, 1982.
- Berchem, J., and C. T. Russell, Flux transfer events on the dayside magnetopause: Spatial distribution and controlling factors, subm. to J. Geophys. Res., 1983.
- Birn, J., and E. W. Hones, Jr., Three-dimensional computer modeling of dynamic reconnection in the geomagnetic tail, J. Geophys. Res., **86**, 6802, 1981.
- Birn, J., and K. Schindler, Two-ribbon flares: magnetospheric equilibria, in Solar Flare Magnetohydrodynamics, (ed. E. R. Priest), p. 337, Gordon and Breach, London, 1981.

- Birn, J., R. Sommer, and K. Schindler, Open and closed magnetospheric tail configurations and their stability, Astrophys. Space Sci., **35**, 389, 1975.
- Cargill, P. J., and E. R. Priest, Slow-shock heating and the Kopp-Pneuman model for 'post'-flare loops, Solar Phys., **76**, 357, 1982.
- Cheng, C.-C., E. C. Bruner, E. Tandberg-Hanssen, B. E. Woodgate, R. A. Shine, P. J. Kenney, W. Henze, and G. Poletto, Observations of solar flare transition zone plasmas from the solar maximum mission, Astrophys. J., **253**, 353, 1982.
- Coppi, B., R. Galvao, R. Pellat, M. N. Rosenbluth, and P. H. Rutherford, Resistive internal kink modes, Sov. J. Plasma Phys., **2**, 533, 1976.
- Coroniti, F. V., and A. Eviatar, Magnetic field reconnection in a collisionless plasma, Ap. J. Supp., **33**, 189, 1977.
- Coroniti, F. V., L. A. Frank, D. J. Williams, R. P. Lepping, F. L. Scarf, S. M. Krimigis, and G. Gloeckler, Variability of plasma sheet dynamics, J. Geophys. Res., **85**, 2957, 1980.
- Cowley, S. W. H., A self-consistent model of a simple magnetic neutral sheet system surrounded by a cold, collisionless plasma, Cosmic Electrodyn., **3**, 488, 1973.
- Cowley, S. W. H., Plasma populations in a simple open model magnetosphere, Space Sci. Rev., **26**, 217, 1980.
- Cowley, S. W. H., The causes of convection in the Earth's magnetosphere: A review of developments during the IMS, Rev. Geophys. Space Phys., **20**, 531, 1982.
- Cowley, S. W. H., and R. Pellat, A note on adiabatic solutions of the one-dimensional current sheet problem, Planet. Space Sci., **27**, 265, 1979.
- Cowley, S. W. H., and D. J. Southwood, Some properties of a steady state geomagnetic tail, Geophys. Res. Lett., **7**, 833, 1980.
- Cowley, S. W. H., and P. Shull, Jr., Current sheet acceleration of ions in the geomagnetic tail and the properties of ion bursts observed at the lunar distance, Planet. Space Sci., **31**, 235, 1983.
- Crooker, N. U., Dayside merging and cusp geometry, J. Geophys. Res., **84**, 951, 1979.
- Daly, P. W., D. J. Williams, C. T. Russell, and E. Keppler, Particle signature of magnetic flux transfer events at the magnetopause, J. Geophys. Res., **86**, 1628, 1981.
- DeCoster, R. J., and L. A. Frank, Observations pertaining to the dynamics of the plasma sheet, J. Geophys. Res., **84**, 5099, 1979.
- Drake, J. F., J. D. Huba, and N. T. Gladd, Magnetic field diffusion and dissipation in reversed field plasmas, Phys. Fluids, **24**, 78, 1981.
- Dubinin, E. M., I. M. Podgorny, and Y. N. Potanin, Structure of the magnetic field at the boundary of the magnetosphere: analysis of a simulation experiment, Cosmic Research, **18**, 77, 1980.
- Dungey, J. W., Conditions for the occurrence of electrical discharges in astrophysical systems, Phil. Mag., **44**, 725, 1953.

- Dungey, J. W., Interplanetary magnetic fields and the auroral zones, Phys. Rev. Lett., **6**, 47, 1961.
- Eastwood, J. W., Consistency of fields and particle motion in the "Speiser" model of the current sheet, Planet. Space Sci., **20**, 1555, 1972.
- Eastwood, J. W., The warm current sheet model and its implications on the behaviour of the geomagnetic tail, Planet. Space Sci., **22**, 1641, 1974.
- Forbes, T. G., and E. R. Priest, Numerical study of line-tied magnetic reconnection, Solar Phys., **81**, 303, 1982a.
- Forbes, T. G., and E. R. Priest, Neutral line motion due to reconnection in two-ribbon solar flares and magnetospheric substorms, Planet. Space Sci., **30**, 1183, 1982b.
- Forbes, T. G., and E. R. Priest, A numerical experiment relevant to line-tied reconnection in two-ribbon flares, Solar Phys., **84**, 169, 1983.
- Forbes, T. G., E. W. Hones, Jr., S. J. Bame, J. R. Asbridge, G. Paschmann, N. Sckopke, and C. T. Russell, Evidence for the tailward retreat of a magnetic neutral line in the magnetotail during substorm recovery, Geophys. Res. Lett., **8**, 261, 1981.
- Formisano, V., A. Pedersen, and P. -A. Lindqvist, The fine structure of the front side magnetopause during two successive crossings, J. Geophys. Res., **87**, 2115, 1982.
- Francfort, P., and R. Pellat, Magnetic merging in collisionless plasmas, Geophys. Res. Lett., **3**, 433, 1976.
- Frisch, U., A. Pouquet, P. Solem, and M. Meneguzzi, The dynamics of two dimensional magnetohydrodynamics, to appear, J. de Mechanique Theorique et Applique, 1983.
- Furth, H. P., J. Killeen, and M. Rosenbluth, Finite-resistivity instabilities of a sheet pinch, Phys. Fluids, **6**, 459, 1963.
- Fyfe, D., D. Montgomery, and G. Joyce, Dissipative forced turbulence in two dimensional magnetohydrodynamics, J. Plasma Phys., **17**, 369, 1977.
- Galeev, A. A., and L. M. Zelenyi, Tearing instability in plasma configuration, Soviet Phys. JETP, **43**, 1113, 1976.
- Galeev, A. A., F. V. Coroniti, and M. Ashour-Abdalla, Explosive tearing-mode reconnection in the magnetospheric tail, Geophys. Res. Lett., **5**, 707, 1978.
- Gekelman, W., R. L. Stenzel, and N. Wild, Magnetic field line reconnection experiments, Physica Scripta, **T2/2**, 277, 1982.
- Giachetti, R., G. Van Hoven, and C. Chiu, The structure of coronal magnetic loops, Solar Phys., **55**, 371, 1977.
- Goldstein, H., and K. Schindler, Large-scale collision-free instability in two-dimensional plasma sheets, Phys. Rev. Lett., **48**, 1468, 1982.
- Golub, L., G. Noci, G. Poletto, and G. S. Vaiana, Active region coronal evolution, Astrophys. J., **259**, 359, 1982.



- Gosling, J. T., J. R. Asbridge, S. J. Bame, W. C. Feldman, G. Paschmann, N. Sckopke, and C. T. Russell, Evidence for quasi-stationary reconnection at the dayside magnetopause, J. Geophys. Res., **87**, 2147, 1982.
- Heyvaerts, J., and E. R. Priest, Coronal heating by phase-mixed shear Alfvén waves, Astron. Astrophys., **117**, 220, 1983.
- Heyvaerts, J., E. R. Priest, and D. M. Rust, An emerging flux model for the solar flare phenomenon, Astrophys. J., **216**, 123, 1977.
- Hill, T. W., Magnetic merging in a collisionless plasma, J. Geophys. Res., **80**, 4689, 1975.
- Hones, E. W., Jr., Plasma flow in the plasma sheet and its relation to substorms, Radio Sci., **8**, 979, 1973.
- Hones, E. W., Jr., Plasma flow in the magnetotail and its implication for substorm theories, in Dynamics of the Magnetosphere, ed. S. -I. Akasofu, p. 545, D. Reidel, Dordrecht, Holland, 1980.
- Hood, A. W., and E. R. Priest, Kink instability of solar coronal loops as the cause of solar flares, Solar Phys., **64**, 303, 1979.
- Hood, A. W., and E. R. Priest, Magnetic instability of coronal arcades as the origin of two-ribbon flares, Solar Phys., **66**, 113, 1980.
- Hood, A. W., and E. R. Priest, Critical conditions for magnetic instabilities in force-free coronal loops, Geophys. Astrophys. Fluid Dynamics., **17**, 297, 1981.
- Howard, R., and B. J. LaBonte, Surface magnetic fields during the solar activity cycle, Solar Phys., **74**, 131, 1981.
- Huba, J. D., N. T. Gladd, and K. Papadopoulos, The lower-hybrid-drift instability as a source of anomalous resistivity for magnetic field line reconnection, Geophys. Res. Lett., **4**, 125, 1977.
- Huba, J. D., J. F. Drake, and N. T. Gladd, Lower-hybrid-drift instability in field reversed plasmas, Phys. Fluids, **23**, 552, 1980.
- Jaeger, E. F., and T. W. Speiser, Energy and pitch angle distributions for auroral ions using the current sheet acceleration model, Astrophys. Space Sci., **28**, 129, 1974.
- Kadomtsev, B. B., Disruptive instability in tokamaks, Sov. J. Plasma Phys., **1**, 389, 1975.
- Kolmogorov, A. N., The structure of turbulence in incompressible viscous fluid for very large Reynolds numbers, C. R. Acad. Sci. U.R.S.S., **30**, 301, 1941.
- Kopp, R., and G. Pneuman, Magnetic reconnection in the corona and the loop prominence phenomenon, Solar Phys., **50**, 85, 1976.
- Kraichnan, R., and D. Montgomery, Two dimensional turbulence, Rep. Prog. Phys., **43**, 547, 1980.

- Krimigis, S. M., and E. T. Sarris, Energetic particle bursts in the Earth's magnetotail, in Dynamics of the Magnetosphere, (ed. S. -I. Akasofu), p. 599, D. Reidel, Dordrecht, Holland, 1980.
- Lee, L. C., Ion two-stream and modified two-stream instabilities in the magnetic neutral sheet, Geophys. Res. Lett., **9**, 1159, 1982.
- Leighton, R. B., Transport of magnetic fields on the sun, Astrophys. J., **140**, 1547, 1964.
- Levine, R. H., Open magnetic fields and the solar cycle, Solar Phys., **79**, 203, 1982.
- Levy, R. H., H. E. Petschek, and G. L. Siscoe, Aerodynamic aspects of the magnetospheric flow, AIAA J., **2**, 2065, 1964.
- Lyons, L. R., and T. W. Speiser, Evidence for current sheet acceleration in the geomagnetic tail, J. Geophys. Res., **87**, 2276, 1982.
- Matthaeus, W. H., Reconnection in two dimensions: localization of vorticity and current near magnetic X-points, Geophys. Res. Lett., **9**, 660, 1982.
- Matthaeus, W. H., and D. Montgomery, Selective decay hypothesis at high mechanical and magnetic Reynolds numbers, Ann. New York Acad. Sci., **357**, 203, 1980.
- Matthaeus, W. H., and D. Montgomery, Nonlinear evolution of the sheet pinch, J. Plasma Phys., **25**, 11, 1981.
- Maxson, C. W., and G. S. Vaiana, Determination of plasma parameters from soft X-ray images for coronal holes (open magnetic field configurations) and coronal large-scale structures (extended closed-field configurations), Astrophys. J., **215**, 919, 1977.
- Meneguzzi, M., U. Frisch, and A. Pouquet, Helical and nonhelical turbulent dynamos, Phys. Rev. Lett., **47**, 1060, 1981.
- Minami, S., and Y. Takeya, Intrusion of magnetic field into the simulative magnetosphere, Preprint, Faculty of Engineering, Osaka City University, Osaka, Japan, 1983.
- Mok, Y., and G. Van Hoven, Resistive magnetic tearing in a finite length pinch, Phys. Fluids, **25**, 636, 1982.
- Moore, R. L., D. L. McKenzie, Z. Svestka, K. G. Widing, S. K. Antiochos, K. P. Dere, H. W. Dodson-Prince, E. Hiei, K. R. Krall, A. S. Krieger, H. E. Mason, R. D. Petrasso, G. W. Pneuman, J. K. Silk, J. A. Vorpahl, and G. L. Withbroe, The thermal X-ray plasma, in Solar Flares, (ed. P. Sturrock), p. 341, Colorado Assoc. Univ. Press, Boulder, 1980.
- Mozer, F. S., R. B. Torbert, U. V. Fablesen, C. -G. Fälthammar, A. Gonfalone, A. Pedersen, and C. T. Russell, Direct observation of a tangential electric field component at the magnetopause, Geophys. Res. Lett., **6**, 305, 1979.
- Nolte, J. T., M. Gerassimenko, A. S. Krieger, R. D. Petrasso, Z. Svestka, and D. G. Wentzel, Do changes in coronal emission structure imply magnetic reconnection?, Solar Phys., **55**, 401, 1977.

- Nolte, J. T., M. Gerassimenko, A. S. Krieger, and C. V. Solodyna, Coronal hole evolution by sudden large scale changes, Solar Phys., **56**, 153, 1978a.
- Nolte, J. T., A. S. Krieger, and C. V. Solodyna, Short term evolution of coronal hole boundaries, Solar Phys., **57**, 129, 1978b.
- Orszag, S., Numerical simulation of incompressible flows within simple boundaries, Stud. Appl. Math., **50**, 293, 1971.
- Oubukoff, A., On the distribution of energy in the spectrum of turbulent flow, C. R. Acad. Sci. U.R.S.S., **32**, 19, 1941.
- Overskei, D., and P. A. Politzer, Plasma turbulence in the vicinity of a magnetic neutral line, Phys. Fluids, **19**, 683, 1976.
- Papadopoulos, K., The role of microturbulence on collisionless reconnection, in Dynamics of the Magnetosphere, (ed. S.-I. Akasofu), D. Reidel Publishing Co., 1979.
- Park, W., D. A. Monticello, R. B. White, K. McGuire, and M. F. Reusch, Three-dimensional resistive studies of high- $\beta$  PDX tokamak discharges, Proc. U.S. -Japan, Workshop on 3-D MHD Studies, Oak Ridge, TN, October 1981.
- Park, W., D. A. Monticello, and R. B. White, Reconnection rates of magnetic fields, Princeton Plasma Physics Lab, Report 2014, (subm. to Phys. Fluids), 1983.
- Parker, E. N., The solar flare phenomenon and the theory of reconnection and annihilation of magnetic fields, Astrophys. J. Suppl. Series, **8**, 177, 1963.
- Paschmann, G., B. U. Ö. Sonnerup, I. Papamastorakis, N. Sckopke, G. Haerendel, S. J. Bame, J. R. Asbridge, J. T. Gosling, C. T. Russell, and R. C. Elphic, Plasma acceleration at the earth's magnetopause: evidence for reconnection, Nature, **282**, 243, 1979.
- Paschmann, G., G. Haerendel, I. Papamastorakis, N. Sckopke, S. J. Bame, J. T. Gosling, and C. T. Russell, Plasma and magnetic field characteristics of magnetic flux transfer events, J. Geophys. Res., **87**, 2159, 1982.
- Pellat, R., About "reconnection" in a collisionless plasma, Space Sci. Rev., **23**, 359, 1979.
- Pellinen, R. J., and W. J. Heikkila, Energization of charged particles to high energies by an induced substorm electric field within the magnetotail, J. Geophys. Res., **83**, 1544, 1978.
- Petschek, H. E., Magnetic field annihilation, in The Physics of Solar Flares, (ed. W. N. Hess), NASA SP-50, p. 425, Washington D. C., 1964.
- Petschek, H. E., The mechanism for reconnection of geomagnetic and interplanetary field lines, in The Solar Wind, (eds. R. J. Mackin, Jr., and M. Neugebauer), Pergamon Press, p. 257, 1966.
- Pneuman, G. W., Two-ribbon flares: 'post'-flare loops, in Solar Flare Magnetohydrodynamics, (ed. E. R. Priest), 379, Gordon and Breach, New York, 1981.
- Podgorny, I. M., E. M. Dubinin, and Y. N. Potanin, The magnetic field on the magnetospheric boundary from laboratory simulation data, Geophys. Res. Lett., **5**, 207, 1978.

- Podgorny, I. M., E. M. Dubinin, and Y. N. Potanin, On magnetic curl in front of the magnetospheric boundary, Geophys. Res. Lett., **7**, 247, 1980.
- Potter, D., Computational Physics, J. Wiley and Sons, New York, 1977.
- Priest, E. R., (ed.) Solar Flare Magnetohydrodynamics, Gordon and Breach, London, 1981.
- Priest, E. R., Theories for simple-loop and two-ribbon solar flares, Fundamentals of Cosmic Physics., **7**, 363, 1982.
- Pu, Z., K. B. Quest, M. G. Kivelson, and C. Tu, Lower-hybrid-drift instability and its associated anomalous resistivity in the neutral sheet of Earth's magnetotail, J. Geophys. Res., **86**, 8919, 1981.
- Rich, F. J., V. M. Vasyliunas, and R. A. Wolf, On the balance of stresses in the plasma sheet, J. Geophys. Res., **77**, 4670, 1972.
- Rijnbeek, R. P., S. W. H. Cowley, D. J. Southwood, and C. T. Russell, Observations of reverse polarity flux transfer events at the Earth's dayside magnetopause, Nature, **300**, 23, 1982.
- Russell, C. T., and R. C. Elphic, Initial ISEE magnetometer results: Magnetopause observations, Space Sci. Rev., **22**, 681, 1978.
- Rutherford, P. H., Nonlinear growth of the tearing mode, Phys. Fluids, **16**, 1903, 1973.
- Sato, T., and T. Hayashi, Externally driven reconnection and a powerful magnetic energy converter, Phys. Fluids, **22**, 1189, 1979.
- Sato, T., T. Hayashi, R. J. Walker, and M. Ashour-Abdalla, Neutral sheet current interruption and field-aligned current generation by three-dimensional driven reconnection, Geophys. Res. Lett., **10**, 221, 1983.
- Schindler, K., A theory of the substorm mechanism, J. Geophys. Res., **79**, 2803, 1974.
- Schnack, D. D., and J. Killeen, Nonlinear saturation of the tearing mode in a reversed field pinch, J. Nuc. Fusion, **19**, 877, 1979.
- Scholer, M., F. M. Ipavich, G. Gloeckler, D. Hovestadt, and B. Klecker, Leakage of magnetospheric ions into the magnetosheath along reconnected field lines at the dayside magnetopause, J. Geophys. Res., **86**, 1299, 1981.
- Scholer, M., D. Hovestadt, F. M. Ipavich, and G. Gloeckler, Energetic protons, alpha particles, and electrons in magnetic flux transfer events, J. Geophys. Res., **87**, 2169, 1982.
- Scudder, J. D., K. W. Ogilvie, and C. T. Russell, The electron plasma signature of flux transfer intervals: Evidence for ongoing, localized, time-dependent reconnection, subm. to J. Geophys. Res., 1983.
- Sheeley, N. R., Jr., J. D. Bohlin, G. E. Brueckner, J. D. Purcell, V. E. Scherrer, and R. Tousey, The reconnection of magnetic field lines in the solar corona, Astrophys. J., **196**, L129, 1975a.
- Sheeley, N. R., Jr., J. D. Bohlin, G. E. Brueckner, J. D. Purcell, V. E. Scherrer, and R. Tousey, XUV observations of coronal magnetic fields, Solar Phys., **40**, 103, 1975b.

- Smith, R. A., A review of double layer simulations, Physica Scripta, **T2/1**, 238, 1982.
- Solodyna, C. V., A. S. Krieger, and J. T. Nolte, Observations of the birth of a small coronal hole, Solar Phys., **54**, 123, 1977.
- Sonnerup, B. U. Ö., Magnetic field reconnexion in a highly conducting incompressible fluid, J. Plasma Phys., **4**, 161, 1970.
- Sonnerup, B. U. Ö., Adiabatic particle orbits in a magnetic null sheet, J. Geophys. Res., **76**, 8211, 1971.
- Sonnerup, B. U. Ö., Transport mechanisms at the magnetopause, in Dynamics of the Magnetosphere, (ed. S.-I. Akasofu), D. Reidel, p. 77, 1979a.
- Sonnerup, B. U. Ö., Magnetic field reconnection, in Solar System Plasma Physics, (ed. Lanzerotti, Kennel, and Parker) Vol. III, p. 45, North Holland, Amsterdam, 1979b.
- Sonnerup, B. U. Ö., and B. G. Ledley, Electromagnetic structure of the magnetopause and boundary layer, in Magnetospheric Boundary Layers (B. Battrock, ed.), p. 401, Paris: ESA, SP-148, 1979.
- Sonnerup, B. U. Ö., G. Paschmann, I. Papamastorakis, N. Sckopke, G. Haerendel, S. J. Bame, J. R. Asbridge, J. T. Gosling, and C. T. Russell, Evidence for magnetic field reconnection at the Earth's magnetopause, J. Geophys. Res., **86**, 10049, 1981.
- Soward, A., Fast magnetic field-line reconnection in a compressible fluid. part 2, skewed field lines, J. Plasma Phys., **28**, 415, 1982.
- Soward, A., and E. R. Priest, Fast magnetic field-line reconnection in a compressible fluid, part 1, coplanar field lines, J. Plasma Phys., **28**, 335, 1982.
- Speiser, T. W., Particle trajectories in model current sheets, 1. Analytical solutions, J. Geophys. Res., **72**, 4219, 1965.
- Speiser, T. W., Particle trajectories in model current sheets, 2. Application to auroras using a geomagnetic tail model, J. Geophys. Res., **72**, 3919, 1967.
- Speiser, T. W., Conductivity without collisions or noise, Planet. Space Sci., **18**, 613, 1970.
- Spicer, D. S., An unstable arch model of a solar flare, Solar Phys., **53**, 305, 1977.
- Spicer, D. S., Loop models of solar flares: revisions and comparisons, Solar Phys., **70**, 149, 1981.
- Spicer, D. S., Magnetic energy storage and conversion in the solar atmosphere, Space Science Reviews, **31**, 351, 1982.
- Spicer, D. S., Magnetic energy storage and the thermal versus non-thermal hard X-ray hypothesis, Adv. Space Res. (Cospar), Vol. 2, No. 11, 135, 1983.
- Steinolfson, R. S., Energetics and the resistive tearing mode: effects of Joule heating and radiation, Phys. Fluids, in press, 1983.

- Steinolfson, R. S., and G. Van Hoven, The growth of the tearing mode: boundary and scaling effects, Phys. Fluids, **26**, 117, 1983a.
- Steinolfson, R. S., and G. Van Hoven, Nonlinear evolution of the resistive tearing mode, Phys. Fluids, submitted, 1983b.
- Steinolfson, R. S., and G. Van Hoven, Radiative tearing: magnetic reconnection on a fast thermal-instability time scale, Astrophys. J., **276**, Jan. 1, 1984.
- Stenzel, R. L., and W. Gekelman, Laboratory experiments on magnetic field line reconnection, in Proceedings of the XV International Conference on Phenomena in Ionized Gases, Invited Papers, 46, Minsk, July 14-18, 1981.
- Stenzel, R. L., R. Williams, R. Aguero, K. Kitazaki, A. Ling, T. McDonald, and J. Spitzer, A novel directional ion energy analyzer, Rev. Sci. Instrum., **53**, 1027, 1982.
- Stenzel, R. L., W. Gekelman, and N. Wild, Magnetic field line reconnection experiments, Part 5, Current disruptions and double layers, J. Geophys. Res., **88**, 4793, 1983a.
- Stenzel, R. L., W. Gekelman, N. Wild, J. M. Urrutia, and D. Whelan, Directional velocity analyzer for measuring electron distribution functions in plasmas, Rev. Sci. Instrum., **54**, 1302, 1983b.
- Stern, D. P., The role of O-type neutral lines in magnetic merging during substorms and solar flares, J. Geophys. Res., **84**, 63, 1979.
- Sturrock, P., (ed.) Solar Flares, Colo. Assoc. Univ. Press, Boulder, Colorado, 1980.
- Svestka, Z., A. S. Krieger, R. C. Chase, and R. Howard, Trans-equatorial loops interconnecting Mcmath regions 12472 and 12474, Solar Phys., **52**, 69, 1977.
- Sweet, P. A., The neutral point theory of solar flares, in Electromagnetic Phenomena in Cosmic Physics, (ed. B. Lehnert), Cambridge U. Press, p. 123, 1958.
- Sykes, A., and J. Wesson, Relaxation instability in tokamaks, Plasma Phys. and Controlled Nuclear Fusion Res., **1**, 237, 1976.
- Tanaka, M., and T. Sato, Simulations of lower-hybrid-drift instability and anomalous resistivity in magnetic neutral sheet, J. Geophys. Res., **86**, 5541, 1981.
- Ugai, M., Dependence of fast magnetic reconnection on electrical resistivity in an isolated current-sheet system, Phys. Fluids, **26**, 1569, 1983.
- Van Hoven, G., Simple-loop flares: magnetic instabilities, in Solar Flare Magnetohydrodynamics, (ed. E. R. Priest), p. 217, Gordon and Breach, London, 1981.
- Van Hoven, G., R. S. Steinolfson, and T. Tachi, Energy dynamics in stressed magnetic fields: the filamentation and flare instabilities, Astrophys. J., **268**, 860, 1983.
- Van Tend, W., and M. Kuperus, The development of coronal electric current systems in active regions and their relation to filaments and flares, Solar Phys., **59**, 115, 1978.

- Vasyliunas, V. M., Theoretical models of magnetic field line merging. 1, Rev. Geophys. and Space Phys., **13**, 303, 1975.
- Vasyliunas, V. M., Upper limit on the electric field along a magnetic 0 line, J. Geophys. Res., **85**, 4616, 1980.
- von Goeler, S., W. Stodiek, and N. Sauthoff, Internal disruption, Phys. Rev. Lett., **33**, 1201, 1974.
- Waddell, B. V., G. L. Jahns, J. D. Callen, and H. R. Hicks, Internal disruptions in tokamaks, Nucl. Fusion, **18**, 4, 1978.
- Waddell, B. V., B. Carreras, H. R. Hicks, and J. A. Holmes, Nonlinear interaction of tearing modes in highly resistive tokamaks, Phys. Fluids, **22**, 896, 1979.
- Wallenhorst, S. G., and R. Howard, On the dissolution of sunspot groups, Solar Phys., **76**, 203, 1982.
- Wallenhorst, S. G., and K. P. Topka, On the disappearance of a small sunspot group, Solar Phys., **81**, 33, 1982.
- Whelan, D. A., and R. L. Stenzel, Electromagnetic wave excitation in a large laboratory beam-plasma system, Phys. Rev. Lett., **47**, 95, 1981.
- Whelan, D. A., and R. L. Stenzel, Nonlinear energy flow in a beam-plasma system, Phys. Rev. Lett., **50**, 1133, 1983.
- White, R. B., D. A. Monticello, M. N. Rosenbluth, and B. V. Waddell, Saturation of the tearing mode, Phys. Fluids, **20**, 800, 1977.
- Wild, N., R. L. Stenzel, and W. Geikelman, Electron temperature measurements using a 12-channel array probe, Rev. Sci. Instrum., **54**, 935, 1983.
- Williams, D. J., Energetic ion beams at the edge of the plasma sheet: ISEE 1 observations plus a simple explanatory model, J. Geophys. Res., **86**, 5507, 1981.
- Winske, D., Current-driven microinstabilities in a neutral sheet, Phys. Fluids, **24**, 1069, 1981.
- Yang, C. -K., and B. U. Ö. Sonnerup, Compressible magnetic field reconnection: a slow wave model, Astrophys. J., **206**, 570, 1976.
- Yeh, T., and W. I. Axford, On the reconnexion of magnetic field lines in conducting fluids, J. Plasma Phys., **4**, 207, 1970.





## CHAPTER 2

### PARTICLE ACCELERATION

#### WORKING GROUP MEMBERS

R. Rosner, Chairman  
*Harvard College*

E. L. Chupp  
*University of New Hampshire*

G. Gloeckler  
*University of Maryland*

D. J. Gorney  
*The Aerospace Corporation*

S. M. Krimigis  
*Johns Hopkins University*

Y. Mok  
*University of California, Irvine*

R. Ramaty  
*NASA/Goddard Space Flight Center*

D. W. Swift  
*University of Alaska*

L. Vlahos  
*University of Maryland*

E. G. Zweibel  
*University of Colorado*

## CHAPTER 2

### PARTICLE ACCELERATION

I.	Introduction	2-3
II.	Phenomenology	2-3
III.	Theoretical Particle Acceleration Mechanisms	2-20
IV.	The Remaining Questions	2-37
V.	References	2-40

## I. INTRODUCTION

Few phenomena span such a wide range of laboratory and astrophysical applications as charged particle acceleration. As succinctly pointed out by Jokipii (1979), the basic problem is deceptively simple, and revolves about the study of the elementary dynamical equation for the gain in kinetic energy of a given charged test particle

$$\delta K.E. = q \int_0^{\delta t} dt \vec{E}[\vec{r}(t), t] \cdot \vec{v}[\vec{r}(t), t] \quad (1)$$

where the electric field  $\vec{E}[\vec{r}(t), t]$  is determined by the ambient plasma and the detailed orbit of the chosen test particle. In actuality, the precise evolution of particle distribution functions (posed as an initial value problem either in the temporal sense - sudden "turn-on" of an acceleration process - or in terms of spatial injection) is an enormously complex issue, one which is only now starting to be fully understood under conditions which represent actual physical conditions in presumed acceleration sites in a reasonably realistic manner.

The subject of particle acceleration in both the solar terrestrial and the more general astrophysical context has been extensively reviewed within the past few years; and furthermore has been the subject of discussion at a number of recent symposia. Thus, there are a substantial number of readily available, published review articles and proceedings, including a review of particle acceleration processes in the solar context by Heyvaerts (1981) and Forman et al. (1984), the Proceedings of the La Jolla Institute Conference on Particle Acceleration Mechanisms in Astrophysics (Arons et al., 1979), and the Proceedings of the Varenna Workshop on Plasma Astrophysics (Guyenne and Levy, 1981). It therefore did not seem appropriate to provide a fully detailed review of the subject of particle acceleration. Instead, we have chosen to focus on a brief summary of our current understanding and recent developments in both theory and observations, and to emphasize both some of the unity of the theoretical approaches which come to the fore when one examines particle acceleration under a wide range of astrophysical circumstances, and to point out the rather significant differences in both diagnostic tools and detailed physical processes which can prevail.

This report is hence broken up into three major sections: the first (Section II) provides a brief phenomenological account of the subject of particle acceleration, with considerable orientation toward the specific physical circumstances under which particle acceleration is known to take place; the second (Section III) details the current status of theoretical studies of particle acceleration; and, finally (in Section IV), we present the outstanding problems which remain when specific mechanisms and the observations are confronted.

## II. PHENOMENOLOGY

In order to define the problems of particle acceleration which theory is to address, it behooves one to begin with the observations. In this context, it is well to remember that observations of nonthermal particles truly span the full range of astrophysical problems: from the Earth's ionosphere and magnetosphere to other planetary magnetospheres (including, most prominently, the Jovian and Saturnian magnetospheres), to the Sun and its interplanetary environs (including the solar wind), to the galaxy proper (including such sources as late-type and pre-main sequence stars, the shocks separating OB stellar winds from the interstellar medium, pulsars, supernova remnants, the galactic center, and possibly binary X-ray sources), to, finally, extragalactic objects

(including jets and active galactic nuclei). Under any of these vastly disparate circumstances, one is in fact led to ask rather similar questions regarding the basic physics:

- What are the physical conditions in the acceleration region?
- What are the characteristics of the ambient source material (which serves to provide the input particles to the accelerator)?
- What are the consequences of the acceleration process: the composition, velocity distribution (including possible anisotropies), and time dependence of the products of the acceleration process?
- What is the source of free energy in the system, e.g., what is the operative driver?
- What are the characteristics of the acceleration mechanism itself: can one, for example, determine the amplitude level of electric field fluctuations, or fix the efficiency with which a given acceleration process operates?

Yet in spite of these similarities, what is particularly striking when one examines the available observational evidence for particle acceleration is the rather gross differences in observational methods and tools which are used to deduce the nonthermal particle distribution functions. In general, these can be distinguished into two distinct approaches: the first seeks to determine particle properties by in situ measurements (as is done in the ionospheric, magnetospheric, and interplanetary domains), while the second seeks to deduce particle properties through the interaction, and the consequent radiative signatures, of accelerated particles with other matter and fields (as is done in solar and cosmic physics). In the former case, it is usually (though not invariably) possible not only to directly sample the particle spectra and composition, but also to determine physical conditions within the presumed acceleration regions; in the second, conditions in the acceleration region must usually be inferred from, for example, a combination of particle spectroscopy, compositional analysis, and timing measurements. This gross classification, of course, reflects a rather basic (and very likely insuperable) difference in the accessibility of the accelerated particles and the acceleration region to diagnostic probing. In this context it is noteworthy that because of the extremely large mean-free paths and gyroradii of high energy extrasolar cosmic rays, in situ measurements of the spectra and composition of such cosmic rays have long been possible, but their source region has remained a focal point of ongoing controversy.

It is also well to remember that theory approaches the problem of particle acceleration in two very different fashions. In the first, the focus is on the acceleration mechanism proper; that is, one asks how particular physical circumstances -- such as prevail in a neutral sheet or within a high Mach number collisionless shock -- determine the output particle spectrum, angular distribution, composition, and the overall acceleration efficiency. In the second instance, one places specific acceleration mechanisms within a larger ("global") context; for example, one considers collisionless shock particle acceleration at the bowshock under both quasi-perpendicular and quasi-parallel conditions as part of an overall model of the bowshock structure. In the latter case, the question to be answered is how the overall system behaves; here one is interested not only in the particle acceleration process in and of itself, but also in its effects on other physical processes (such as, for example, generation of waves in the foreshock region). This distinction between "process-oriented" and "system-oriented" studies is particularly relevant to the solar-terrestrial domain because the complex system under study is highly interactive; thus, in many cases it is far from straightforward to separate the contributions to observed particle distributions from the various possible contributing processes, and hence not trivial to associate a unique observational signature with specific acceleration mechanisms on the basis of the observations.

## SOLAR FLARES

The properties of energetic particles accelerated during the energy release in solar flares are deduced primarily by: (a) remote sensing of the radiation signatures resulting from the interaction between fast particles and the ambient solar atmosphere, and (b) direct measurements of particle properties in space. The data now available to the solar physicist are principally derived from space-based observations from ATM, the OSO and HEAO series of satellites, SMM, Hinotori, P78-1 and ISEE-3 and ground-based observations in the radio; reviews summarizing these data can be found in Ramaty et al. (1980), Chupp (1982), Kane et al. (1980), and Kundu and Gergely (1980). The data principally consist of hard X-ray (10-400 keV), radio, gamma-ray, and neutron emissions directly from the flare, and electron, proton, and heavy ion energy spectrum and composition measurements obtained in interplanetary space. From the existing data, we can deduce information about the nature of the acceleration process(es), and location of the radiation and acceleration sites(s). In the following, we briefly outline the properties of the solar accelerator as they are known today.

Hard X-rays result from the bremsstrahlung emission from energetic electrons with energies ranging from  $\lesssim 10$  to hundreds of keV with the solar chromosphere. Interpretation of the hard X-ray photons is usually cast in terms of the so-called thin or thick target models (Brown, 1971, 1975), although the thick target model is presently favored. The microwave emission is interpreted as gyro-synchrotron emission from mildly relativistic electrons with energies 100 keV or higher (Ramaty, 1969); this microwave emission is closely associated in time with the hard X-ray emission, and is probably emitted from the acceleration site or its immediate neighborhood. Although the interpretation of the observed spectra is always model dependent, the existing results fit well a power law energy distribution extending at times to energies as low as 10 keV and as high as  $\approx 400$  keV.

With the advent of the SMM and Hinotori satellites, and the onset of observations with the Very Large Array telescope, it has become possible to obtain spatially resolved two-dimensional images of the radiating source(s) in the radio band and at hard X-ray energies less than 30 keV with time resolution of a few seconds. These observations show that while the radio emission is sometimes confined to the upper part of a closed magnetic topology (solar loop; Marsh and Hurford, 1980; Kundu, et al., 1982), the hard X-ray emission mainly comes from the foot points of presumably closed magnetic topologies (Duijveman et al., 1982; Hoyng et al., 1983). These results are consistent with recent stereoscopic observations of high energy ( $< 100$  keV) photon emission from five solar flares and have been made with the X-ray spectrometers aboard the ISEE-3 and Pioneer Venus Orbiter spacecraft (Kane, 1983). The observed altitude structure of the photon source has shown that the impulsive X-ray emission is located at an altitude  $\lesssim 2,500$  km above the solar photosphere. The temporal structure of the impulsive emission is "spikey", and is usually characterized by pulses of short duration (Kiplinger et al., 1982; Kaufmann et al., 1983). In many instances, the hard X-ray bursts are closely associated with type III bursts (Kane et al., 1982), which are the signature of fast streams of nonthermal electrons escaping from the corona. Examination of the starting frequencies of type III burst has also suggested that the injection or acceleration of electrons on open field lines occurs in the low corona (i.e., plasma densities  $\sim 10^9$  cm $^{-3}$ ; Kane et al., 1982). These considerations have led to several suggestions for the physical geometry of the flare site, one of which is illustrated in Figure 2-1: the conjecture is that electrons are accelerated at the interface of two coronal loops. Some of these electrons are trapped within the loops, and radiate at microwave and meter wavelength frequencies. Other electrons follow the open field lines, and either produce associated type III bursts or are directly detected in space; and still others stream towards the chromosphere, where they radiate hard X-ray photons from their interaction with the ions at the transition region or below.

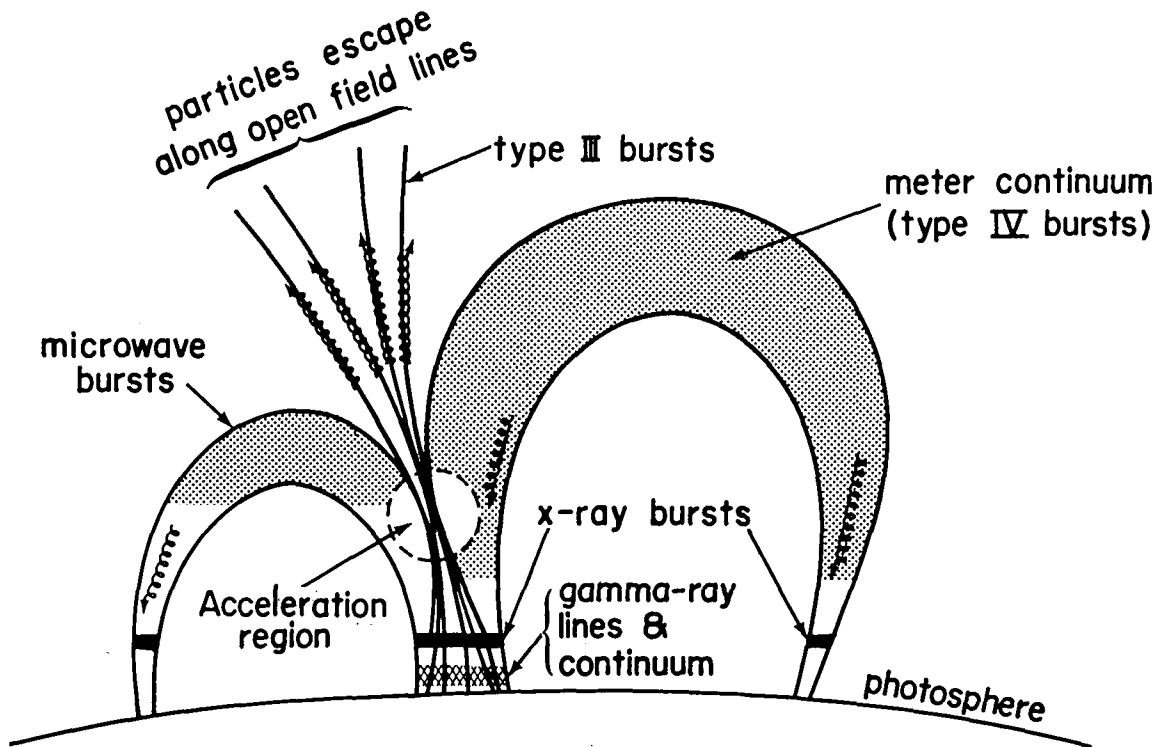


Figure 2-1. A schematized picture of the interaction between two solar coronal magnetic "loops", which defines one of the "standard" scenarios for particle acceleration during the course of a solar flare.

It has been known for some time that, following the impulsive injection of electrons and ions, a shock is developed which accelerates electrons and ions to even higher energies; this shock is responsible for the type II radio emission. This process, whenever it appears, has been associated with second phase acceleration, which had been conjectured as a necessary element in the acceleration of electrons ( $>100$  keV) and ions. As we will discuss below, SMM recent observations of  $\gamma$ -ray lines and continuum have shown that, in many instances, protons and electrons are accelerated nearly simultaneously at all energies at the beginning of the impulsive emission (Chupp, 1982; Forrest and Chupp, 1983).

Observations of  $\gamma$ -rays and high-energy neutrons are a relatively new and useful diagnostic of relativistic particle acceleration in solar flares.  $\gamma$ -ray lines are the product of nuclear reactions between flare-accelerated protons and nuclei with the ambient solar atmosphere. Narrow line emission results from the deexcitation of nuclear levels in solar atmospheric nuclei, such as C, N, O, Ne, Mg, Si and Fe, and from neutron capture and positron annihilation. Broad-band nuclear emission results from the superposition of such lines and deexcitation radiation from excited heavy nuclei in the energetic particle population. Of particular significance is the (4-7) MeV band, in which the nuclear lines of C, N, and O produce the bulk of the observed emission. The strongest narrow line from disk flares is the 2.223 MeV line, resulting from neutron capture on H in the photosphere. The underlying  $\gamma$ -ray continuum in solar flares is produced by bremsstrahlung from relativistic electrons and by broad nuclear emission from the processes mentioned above. Neutrons of energies greater than about 50 MeV produced by the nuclear reactions can survive transit from the Sun to the Earth, and be directly detected at Earth. For flares whose duration is much shorter than the transit time, the time dependence of the observed high energy neutron flux at Earth is a direct measure of the energy spectrum of the neutrons released at the

Sun (Chupp et al., 1982). Theoretical calculations can relate this spectrum to the spectrum of the accelerated particles, thereby providing information on the particle spectrum in the energy range from 100 to 1000 MeV. The neutron observations are thus complementary to the gamma-ray observations. By studying the gamma-ray and neutron observations, we can reach the following conclusions for the acceleration of particles with energies above an MeV:

(a) The spectrum of the energetic particles must not vary much from flare to flare. The proton spectrum cannot be fit by a single power law over the entire range from 10 to 1,000 MeV, but an exponential or Bessel function distribution provides a good fit (a power law distribution is adequate for energies  $\geq 500$  MeV). The spectrum of electrons within the energy range required to produce the gamma-ray continuum can be fit to a power law; the relative number of these electrons above a few MeV is  $10^{-3}$  of the protons above the same energy (Ramaty et al., 1983a, 1983b).

(b) We can bracket the densities (and the height) of the ambient medium at which  $\gamma$ -rays and neutrons are produced; the existing observations suggest that the ambient density is  $10^{12} < n < 10^{16} \text{ cm}^{-3}$ . The acceleration site, on the other hand, is in the low corona (i.e., at much lower densities). Both of these results are in good agreement with the radio and hard X-ray emissions discussed above (Ramaty et al., 1983b).

(c) Studies of the time structure of several  $\gamma$ -ray lines, the  $\gamma$ -ray continuum, and hard X-ray emission suggest that in many events the interaction of electrons and ions with the ambient plasma (and presumably their acceleration) occurs nearly simultaneously, i.e., within seconds. In some events, however, a relative delay of the peak of emission of the high-energy photons ( $>300$  keV) compared to that of the hard X-ray emission is observed. At times, several spikes of  $\gamma$ -ray continuum have been observed, with rise times of a few seconds (Chupp et al., 1981; Forrest et al., 1981).

(d) Comparison of  $\gamma$ -ray events with other emission have shown that there is an excellent correlation with the "brilliance" measure of intensity for H $\alpha$ -associated flares, an excellent correlation with microwave-bursts (at least for the 2-year period of SMM observations), and a fair correlation with the meter wavelength radio emission. The quantitative correlation with solar cosmic rays in space is poor (Chupp and Forrest, 1983).

In situ measurements in space can provide information about the particle spectra, the ratio of electrons and protons in the same energy range, and the composition of the accelerated particles. We can summarize the main conclusions as follows:

(a) A recent study by McGuire et al. (1981) showed that the spectrum of protons and  $\alpha$  particles from 1 to 400 MeV fits well to an exponential or Bessel function distribution.

(b) The number of ratio of electrons and ions observed in space in the same energy range varies from  $10^{-6}$  to  $\approx 1$  (Evenson et al., 1981). The apparent inconsistency with the number of accelerated electrons and protons estimated from  $\gamma$ -ray lines and continuum is thought to be simply a consequence of varying escape efficiency for particles propagating in coronal regions with "closed" (re-entrant) or "open" (to the interplanetary medium) magnetic field lines; the precise mechanism leading to variable escape efficiency is, however, not as yet definitively established.

In recent years a better understanding of the elemental, charge-state, and isotopic composition of energetic ions ( $<1$  to  $>20$  MeV) in solar energetic particle events (SEP) has been achieved (see Fan et al., 1983). The composition within individual SEP events may vary in time and in energy, and will in general be different from that in other SEP events. It has been shown that the average value of relative abundances measured in a large number of SEP events show a systematic deviation from photospheric abundances. The broad spectrum of observed elemental abundance

variations, which in their extreme result in composition anomalies characteristic of  $^3\text{He}$ -rich, heavy-ion rich and carbon-poor SEP events, along with direct measurements of the ionization states of SEPs, suggests that two processes may be involved, e.g., a preferential preheating of the ambient plasma that may include pockets of cold material, followed by a bulk acceleration to the observed energies. We also expect that interplanetary propagation, as well as acceleration by interplanetary propagating shocks, will often further modify the composition of SEP events at lower energies.

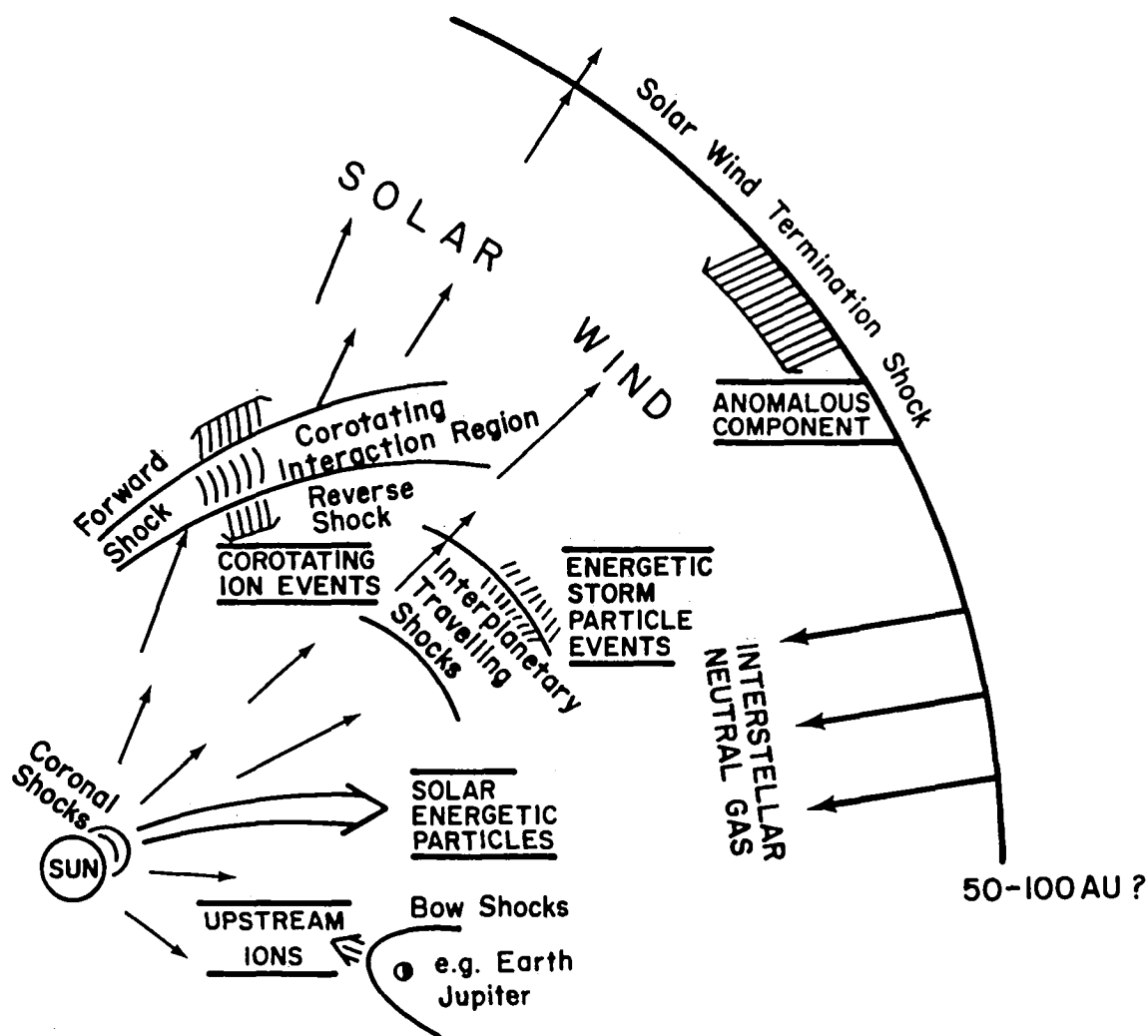
In summary electrons and ions are accelerated in the vicinity of the energy release region in solar flares, and propagate from there into the ambient solar atmosphere and interplanetary medium, where we either observe the results of the interaction between accelerated particles and the ambient plasma, or directly sense their presence. One of the principal challenges of understanding particle acceleration at the Sun is to connect this propagation problem with models for solar magnetic fields, whose topology seems to be a crucial determinant of particle propagation and storage (cf., Zweibel and Haber, 1983, and references therein).

## ACCELERATION PROCESSES IN THE INTERPLANETARY MEDIUM

A variety of acceleration processes have been observed or postulated to exist in the interplanetary medium; nevertheless, some form of shock acceleration appears to provide the most promise in accounting for the observations, as will be discussed below. The likely locations of acceleration regions in the heliosphere, along with the energetic particle populations produced at these sites are shown schematically (not to scale) in Figure 2-2. The acceleration region in each case is associated with shocks which are either remote from Earth (i.e., coronal shocks, forward-reverse shocks in corotating interaction regions, and the solar wind terminal shock), travel outward in the heliosphere (solar flare-associated shocks), or are associated with planetary magnetospheres (e.g., bowshocks). It is believed that the scale size of the shock and its configuration determine the maximum energy of the accelerated ions, while the duration of the intensity increase (observed, say, near Earth) is governed more by interplanetary propagation processes and the magnetic field topology between the acceleration region and the point of observation. Some typical values for both the maximum energy and time scale for the various acceleration phenomena are given in Table 2-1).

Most of the experimental evidence characterizing these acceleration processes (e.g., distribution functions, compositions, temporal evolutions of intensities, anisotropies) comes from satellite observations near Earth. Not until the measurements of intensity gradients with instrumented deep-space probes (Pioneer 10, 11, Voyager 1, 2) was it possible, however, to identify and locate the remote acceleration sites beyond the orbit of Earth. As new regions of the heliosphere are probed -- out-of-ecliptic with the Solar Polar Mission, and the distant heliosphere with Pioneers and Voyagers -- and as the previously inaccessible suprathermal energy range ( $\sim 10$  to  $\sim 100$  keV) is explored in detail, we will surely gain more insight into the various acceleration processes and likely uncover additional heliospheric acceleration phenomena. Our current understanding of particle acceleration associated with: (a) interplanetary traveling shocks -- Energetic Storm Particles (ES) and Shock Spikes; (b) corotating interactions regions -- Corotating Particle Streams; (c) shocks in the distant heliosphere (solar wind termination shock?) -- Anomalous Cosmic Ray component; and (d) the Earth's bowshock -- Diffuse Upstream Ions, is described briefly below. Low energy solar energetic particles have been discussed in the previous section and further aspects of planetary bowshock associated ions are given in the Section entitled "Bowshock". Recent review articles on various aspects of these topics include McDonald (1981), Lee and Fisk (1982), and Gloeckler (1979).





## PARTICLE ACCELERATION IN THE HELIOSPHERE

Figure 2-2. Schematic representation of acceleration processes observed in the interplanetary medium, showing the likely acceleration sites.

Table 2-1.  
Interplanetary Acceleration Phenomena

Phenomena	Energy	Duration	Seed Population	Acceleration		
				Region	Configuration <sup>†</sup>	Process
1. Solar Energetic Particles (SEPs)	$\lesssim 0.1$ to $\sim 10$ MeV/n	Days	$\sim 10^6$ K Corona + cold material	Corona	$Q_{\parallel}/Q_{\perp}$	Diffusive compression/shock drift
2. Anomalous Cosmic Rays	$\sim 1$ to 30 MeV/n	Years <sup>*</sup>	Interstellar gas?	Heliospheric boundary or Galaxy?	$Q_{\parallel}/Q_{\perp}$ ?	" ?
3. Corotating Ion Streams	$\sim 0.2$ to 10 MeV/n	Days <sup>*</sup>	Solar Wind?	F-R Shocks ( $\gtrsim 2$ AU)	$Q_{\parallel}/Q_{\perp}$	"
4. a) Energetic Storm Particles (ESPs)	$\sim 10$ 's keV to few MeV	Hours-day	Solar Wind?	Propagating	$Q_{\parallel}/Q_{\perp}$	"
b) Shock Spikes	$\sim 10$ 's keV to few MeV	Minutes-hours	Solar Wind?	Interplanetary Shock	$Q_{\perp}$	Shock drift (scatter-free)
5. Diffuse Upstream Ions (Post-shock Spikes)	$\lesssim 10$ keV to $\sim 299$ keV/charge	Hours	Solar Wind?	Planetary Bow Shock or Magnetosphere	$Q_{\parallel}/Q_{\perp}$ ?	Diffuse compression/shock drift ?

\* Observed primarily near solar minimum.

†  $Q_{\perp}$  = Quasi-perpendicular shock;  $Q_{\parallel}$  = Quasi-parallel shock.

## Interplanetary Shock Acceleration

Energetic Storm Particle (ESP) and Shock Spike enhancements associated with propagating interplanetary shocks have been studied extensively both experimentally (see, e.g., review by McDonald, 1981) and theoretically (see, e.g., review by Lee, 1983) over the past 10 to 15 years. ESP events are often associated with interplanetary shocks generated by solar flares. One typically observes a large increase ( $\sim 10$  to  $>100$  fold) in intensity of protons, helium and heavier ions at energies mostly below 1 MeV/nucleon lasting for days (e.g., Klecker et al., 1981; Scholer et al., 1983). The intensity peaks at the shock and falls-off exponentially with distance ahead of the shock, with the e-folding distance ahead of the shock being approximately  $X^* = 10^{13} T^{\alpha}$  cm, where  $\alpha \sim 1$  and  $T$  is the kinetic energy of units of MeV (Klecker et al., 1981). The increase behind the shock is often terminated by a tangential discontinuity in the magnetic field, presumably the boundary between the shocked solar wind and the driver gas. While the energy spectra above several hundred keV are exponential in momentum (rigidity), at lower energy both the spectral shapes and their evolution are more complex (e.g., Gosling et al., 1981; Wenzel et al., 1981); still, near the shock, the spectra often exhibit power-law behavior. The flow direction of ions below a few hundred keV change from streaming away from the shock prior to shock passage to a broad distribution of pitch angles centered perpendicular to the  $\mathbf{B}$  field; however, isotropic distributions have also been observed. ESP increases have been observed not only near Earth but also inside 1 AU with the Helios spacecraft, and at distances  $>20$  AU with Pioneer 10 and 11 (although only limited studies of ESP events beyond 10 AU are available at this time).

Shock spike events (see, e.g., Armstrong et al., 1977) are large (factor of 10 to >100), short duration (~0.1 to 1 hours) increases of ion intensities often seen at or near the passage of interplanetary shocks. The energy spectra are steep, although at times ions with energy in excess of 10 MeV have been observed. High energy electrons (>10 keV) are rarely seen in association with interplanetary shocks (Armstrong and Krimigis 1976, Fillius et al., 1983), although measurements of electrons of a few keV are common.

The observations described above are qualitatively consistent with diffusive shock acceleration (nearly isotropic ESP events, >100 keV ESP ions) and with scatter-free shock reflection (shock spikes). Less certain and still unresolved is the question of the seed population for these energetic ions, although the solar wind is a strong candidate (Gosling et al., 1981). It would, however, be surprising if some of the ambient low-energy solar-flare particles were not also accelerated to higher energy by these processes. One sees evidence for this in post-shock spikes (similar to shock spikes in appearance, but occurring hours after passage of the shock front; Gloeckler et al., 1974), where ambient energetic ions (ESP ions in this case) ahead of the interplanetary shock are further accelerated by interaction with the Earth's bowshock (Scholer and Ipavich, 1983; also see Section on "Ions Upstream of Planetary Bowshocks" below).

Progress on the question of the seed population should come from detailed studies of compositions in both the solar wind and the accelerated ions. Ideally, one needs to measure the time evolution of the entire distribution function for several ion species to settle this question. Composition studies with instruments on the ISEE-1 and -3 spacecraft similar to those done in diffuse upstream events (see Section on "Ions Upstream of Planetary Bowshocks" below) should provide important information on this question. However, the detailed comparison with the solar wind and energetic ion compositions cannot be done until the launch of the better-instrumented ISPM spacecraft in 1986.

### Corotating Ion Streams

Modest increases in the intensity of  $<1$  to  $\sim 20$  MeV/n  $H^+$  and  $He^{++}$  which persisted for 3 to 6 days and often reappeared at 27-day intervals were first observed near Earth during the 1965 solar minimum activity period. A decade later, it was established using Pioneer 10 and 11 instruments (McDonald et al., 1976; Barnes and Simpson, 1976; also see Tsurutani et al., 1982 and references therein) that the acceleration of these corotating ion streams takes place in the vicinity of forward and reverse shocks which form beyond a few AU at the edges of the interaction regions between long-lived solar wind stream. Corotating ions are seen at Earth when the acceleration region and Earth become field-lined connected. Because long-lived solar wind streams disappeared with the onset of increased solar activity, the most recent observations on radial gradients, anisotropies, spectral shapes and composition date back several years and may be summarized as follows (see reviews by Gloeckler, 1979 and McDonald, 1981). The peak intensity of these corotating events reached their highest levels between about 2 and 4 AU, and near Earth these ions were seen to flow toward the Sun (Marshall and Stone, 1977; Mewaldt et al., 1978). The average relative abundances were found to resemble more closely those of the solar wind (or corona) than those of solar energetic particles, in particular for H, He, C, Si and Fe. The distribution functions of the major ion species in corotating ion events observed near Earth were found to have the form  $f=f_0 \exp(-v/v_0)$  in the range 0.16 to 10 MeV, and the characteristic speed  $v_0$  (typically  $3 \times 10^8$  cm/sec) was determined to be about the same for all these major ion species. These results provide evidence for acceleration at the forward-reverse shocks at distances  $\sim 2$  to  $\sim 4$  AU of the suprathermal distribution of the solar wind.

Although analysis of existing interplanetary data from the Voyager spacecraft could provide some additional information, further progress in understanding the nature of this interplanetary acceleration mechanism and in identifying the seed population will come from observations with ion composition instruments covering a broad energy range during the in-ecliptic path to Jupiter by

ISPM and Galileo in the 1986-87 solar minimum. Interesting studies of corotating particle streams near Earth (e.g., detailed elemental composition measurements with high mass resolution) which would provide important information on the seed population will require large geometry factor instruments.

### The Anomalous Cosmic Ray Component

Possibly the least understood among the large-scale interplanetary acceleration processes is the mechanism which produces the Anomalous Cosmic Ray Component present primarily during solar minimum (and possibly during alternate solar cycles). In fact, it is not entirely clear that this component is heliospheric (i.e., accelerated in the outer heliosphere or at its boundary), rather than galactic in origin. The Anomalous Component appears as a broad hump centered at about 50-70 MeV in the spectra of quiet-time (absence of solar particles) cosmic ray He, N, O, and Ne (see review by Gloeckler, 1979; also Webber and Cummings, 1983). Because this component has a positive radial gradient and is modulated in a fashion similar to the galactic cosmic rays, its origin must be in the distant heliosphere, probably beyond  $\sim 20$  AU. In the heliospheric origin model of Fisk et al. (1974), the Anomalous Component results from the acceleration of interstellar neutral gas (H, He, N, O and Ne) swept into the heliosphere by the motion of the solar system through the interstellar medium. Upon reaching the inner heliosphere, this gas is ionized, carried outward with the solar wind, and accelerated (possibly at the solar wind termination shock). A firm prediction of this model is that the  $\sim 50$ -70 MeV particles are singly-ionized. Thus, decisive to the question of origin (galactic vs. heliospheric) will be direct measurements of the charge states of e.g., the anomalous oxygen. Such observations will unfortunately not be made in the foreseeable future because of a lack of adequate instrumentation and launch opportunities. Observations of the energy spectra and composition (including isotopes) could also help in differentiating between a heliospheric vs. galactic origin. Although high-mass resolution, large area instruments are best suited for these measurements (none of which are currently planned), data from the ISPM and Voyager cosmic ray experiments are likely to provide the best compositional and spectral information on this component in the foreseeable future.

### Ions Upstream of Planetary Bowshocks

A variety of energetic particle populations are observed upstream of the Earth's, Jupiter's, and Saturn's bowshocks; although their acceleration is properly discussed under the magnetospheric rubric, we will summarize their basic phenomenological properties -- as seen in the upstream interplanetary medium -- here. It is noteworthy that the acceleration of diffuse (i.e., isotropic in the spacecraft frame) ions observed upstream of the Earth's bowshock is probably the best diagnosed and understood particle acceleration process (a discussion of the nature of the diffuse (and reflected) components is given below in the Section on "Bowshock". Observations of: (a) inverse velocity dispersion, (b) quasi-exponential distribution functions of 30 to  $\sim 130$  keV/charge  $H^+$ ,  $He^{++}$ , and  $Z > 2$  ions with comparable e-folding energies of typically  $\sim 20$  keV/charge, and (c) an exponential fall-off in intensity and a systematic variation in the anisotropy with distance from the bowshock for those upstream events for which the magnetic field is radial are well explained by a Fermi-type process wherein ions in the reflected (highly directional,  $\sim$  few keV) component are accelerated through multiple scattering between the turbulence near the bowshock and in the upstream region (Terasawa, 1979, and references therein; Lee, 1982).

Current experimental studies include detailed correlations of the diffuse ion intensity with solar wind density ( $n$ ), the time of field-line connection to the bowshock (TOC), the distance from the bowshock ( $L$ ), and the various shock Mach numbers. Preliminary results indicate that there is good correlation of the intensity of 30 keV ions with  $n$  and  $L$ , while the intensity of 130 keV particles correlates well with TOC and Mach number. That the most likely seed population for

the diffuse component is ultimately the solar wind follows from observations of the composition of upstream ions and comparison with the solar wind composition. It is found that there is correlation between the solar wind  $\text{He}^{++}/\text{H}^+$  ratio, and this ratio in the diffuse ion events and that the  $\text{He}^{++}$  is about 50% more abundant in the accelerated ions. This overabundance is also explained by current acceleration models. Although no one-to-one comparison of the heavy ion composition in the upstream diffuse ions with that in the solar wind can as yet be made, it is observed that the distribution of charge-states and the elemental abundances in the diffuse component are similar to those of the typical solar wind. The coronal temperatures inferred from the charge-state distributions of the energetic ions are consistent with those expected for the various types of solar wind flows observed during the occurrence of upstream events.

The question of exactly how some solar wind ions are transformed into the reflected component which forms the seed population for the diffuse ions is still not settled. The reflected component is either produced by solar wind ion reflection at the bowshock (Schwartz et al., 1983) or by leakage of shock-processed ions back upstream (Tanaka et al., 1983). Although analysis of current experimental evidence presumes energization upon reflection or heating at the bowshock, the possibility exists that some solar wind ions are first incorporated and energized in the magnetosphere, and that these energetic magnetospheric ions may then leak out and, with possible additional acceleration upstream, be observed as diffuse ions (see Section on "Bowshocks" below). (Observations of energetic ion increases upstream of the Jovian bowshock show the presence of sulfur which has a clear Jovian origin). This question is presently being studied, but additional observations from better instrumented spacecraft (e.g., AMPTE-IRM for the Earth's bowshock and ISPM and Galileo for the Jovian bowshock) may be required for its final resolution.

## MAGNETOSPHERIC AND IONOSPHERIC OBSERVATIONS

The Earth's magnetosphere, as well as the magnetospheres of the major planets (particularly those of Jupiter and Saturn), are remarkably rich in phenomena leading to the production of fast (nonthermal) particles. Because of the rather distinct physical processes involved, it is useful to segregate this portion of the discussion into three parts: that dealing with auroral phenomena, that dealing with magnetotail acceleration, and, finally, that dealing with particle acceleration at the bowshock. Throughout, it should be kept in mind that although the present discussion concerns itself virtually exclusively with the terrestrial magnetosphere, analogous studies have been conducted in the magnetospheres of other planets (principally Saturn and Jupiter); hence, the terrestrial magnetosphere can be regarded as just one example of a class of physical systems accessible to direct sensing measurements, and the risks inherent in interpreting phenomena unique to a given physical system are largely avoided.

### Auroral Particle Acceleration

Auroral formations are often described as discrete or diffuse, depending morphologically on latitudinal scale lengths or physically on whether low altitude particle acceleration is involved. The diffuse aurora is caused by the precipitation of energetic plasma (ions and electrons) which has been energized in the Earth's magnetotail, and this acceleration process is described in a separate section. The particle acceleration responsible for discrete auroral events occurs in the Earth's inner magnetosphere, where electrons are accelerated downward and ions are accelerated upward along magnetic field lines. These events take on the classic "inverted-V" configuration (Frank and Ackerson, 1971) when viewed with energetic particle detectors at low (<10,000 km) altitudes.

Figure 2-3 illustrates schematically some of the physical features of this acceleration region inferred from field and particle observations by a satellite at an altitude of one Earth radius ( $1 R_e$ ) -- a typical altitude of the S3-3 satellite, for example. The satellite is shown moving from the polar regions toward lower latitude. The roughly V-shaped heavy lines represent electric equipotential contours, indicating a potential gradient along magnetic field lines over a limited latitude range. The figure shows that at some latitudes the potential drop along the magnetic field is predominantly above the satellite; at others, it is predominantly below the satellite; and that there are regions of transition, in which potential drops exist both above and below the satellite. Note that the potential structure is shown to extend for large distances (thousands of kilometers) along the magnetic field, emphasized by the fact that satellites in this altitude regime usually fly through the auroral acceleration region rather than above or below it.

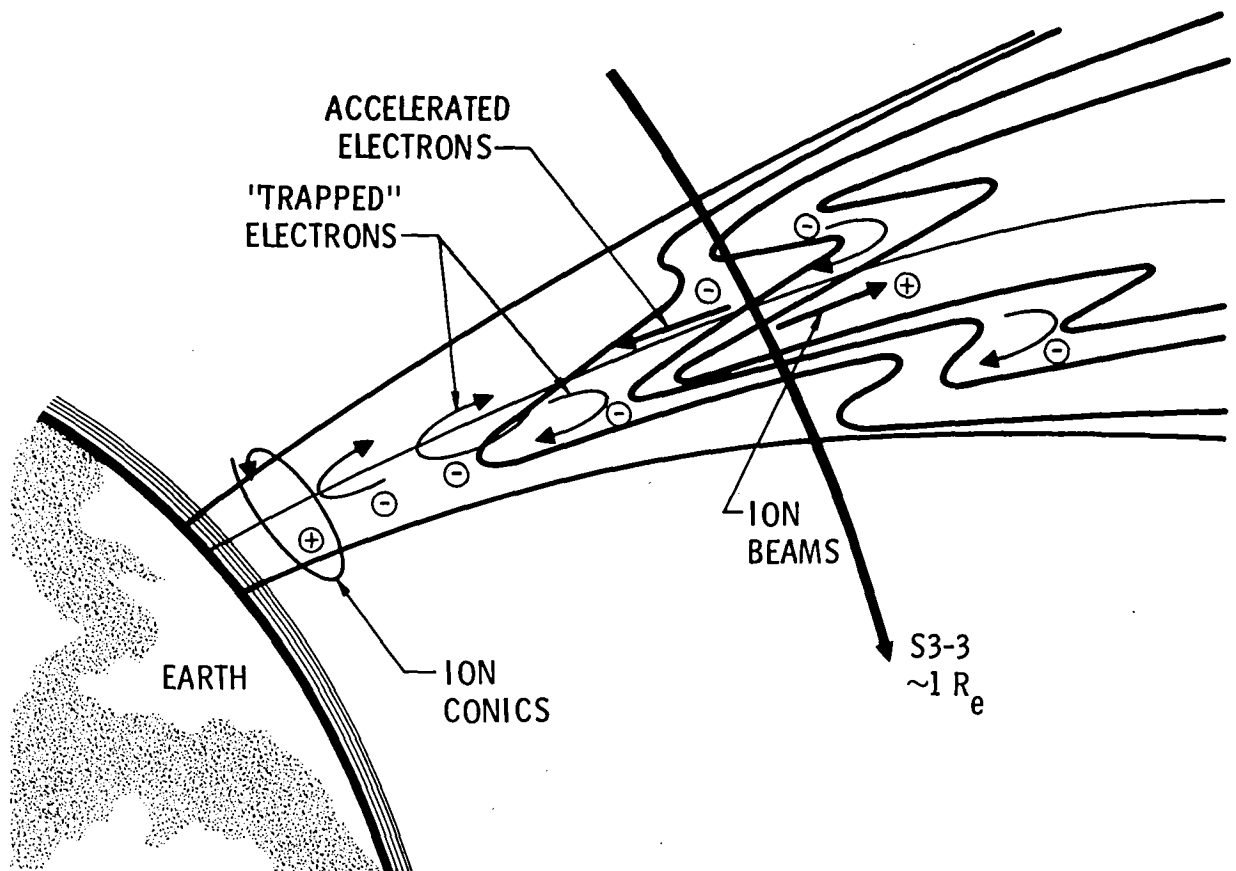


Figure 2-3. A schematic summary of important plasma populations encountered by a low altitude satellite in the auroral region. The dark wavy lines represent probable electric equipotentials, which do not follow the dipole-like magnetic field lines, thus indicating parallel potential drops. The main plasma signatures are upflowing accelerated ions, downflowing accelerated electrons, and trapped electrons. Note the V-shaped potentials, which are thought to be germane to the generation of perpendicular structures in aurorae.

The strong electric fields perpendicular to the magnetic field implied by the illustration have been observed (see Mozer et al., 1980 and references therein), as have parallel electric fields. Perpendicular electric fields as large as several hundred millivolts per meter have been observed on spatial scales of tens of kilometers. The parallel component of the electric field has been observed, with electric fields of about 10 mV/m and localized potential drops on the order of a volt (i.e.,  $e\delta\phi \lesssim T_e$ ; Temerin et al., 1982). These were interpreted as micro-double layers by Temerin et al., but it must be noted that (strong turbulence) anomalous resistivity produced by finite amplitude ion-cyclotron waves can also lead to "spikey" parallel electric fields (Rowland and Palmadesso, 1983). Indeed, since large-amplitude ion-cyclotron waves were observed at the same time as the "spikey" parallel fields, the precise mechanism leading to parallel d.c. electric fields is still unclear (cf., Section on "Double Layers" below). Furthermore, single spacecraft observations are insufficient to determine the actual distribution of double layers in space and, therefore, some question remains concerning the role of micro-double layers in accounting for all observed auroral acceleration, particularly in events exceeding several keV energies.

A number of particle populations result from the electric and magnetic configuration shown in Figure 2-3. They include a downward-accelerated magnetospheric electron population, an electron population trapped between magnetic mirror points at low altitude and an electrostatic potential barrier at high altitude, and an upward-accelerated ion population of ionospheric origin (see reviews by Johnson, 1979; Shelley, 1979; Young, 1979; and Chiu et al., 1983). Each of these energetic particle populations have been observed (Mizera et al., 1981a, b), and the particle observations have been used in turn to infer the extent of the acceleration region in altitude (Ghielmetti et al., 1978; Gorney et al., 1981). It is found that the auroral acceleration region extends to altitudes of about  $3 R_E$ , but at times can be confined below 10,000 km altitude. Total potential drops of over 30 kilovolts have been observed, with several kilovolts being most common.

Although both ions and electrons can exhibit substantially adiabatic behavior in these acceleration regions, a number of distinctly non-adiabatic features have been observed. Upflowing ion beams, for example, typically have temperatures far in excess of the ionospheric source plasma, and mass-dependent energization has been observed (Collin et al., 1981). Another example of non-adiabatic processes is the observation of strongly magnetic field-aligned electron fluxes over a wide energy range of low altitudes, particularly near the edges of acceleration regions (Whalen and Daly, 1977; Craven and Frank, 1975; Fennell et al., 1979). Usually occurring in spatially limited regions, these high fluxes often cause a localized brightening in an otherwise extended auroral form. Field-aligned electron fluxes over a wide energy range cannot be the direct consequence of parallel electrostatic acceleration, but may be the result of energy scattering of an initial mono-energetic beam or perhaps cross-field plasma diffusion into the acceleration region.

Another accelerated particle population (ion conics) is often observed on or near auroral magnetic field lines, and is indicated schematically in Figure 2-3 as having a low altitude ( $\sim 2000$  km) source. Ion conics have high perpendicular temperatures (up to several hundred eV) and have been observed well away from regions of parallel particle acceleration as well as at low altitudes below regions of parallel acceleration (Shelley et al., 1976a, b; Sharp et al., 1977; Klumpp, 1979; Gorney et al., 1981). Possible mechanisms for the production of ion conics include electrostatic ion cyclotron stochastic acceleration (Papadopoulos et al., 1980) or heating (Ashour-Abdalla et al., 1981; Dusenberry and Lyons, 1981; Lysak et al., 1980), lower hybrid wave heating (Chang and Coppi, 1981) and acceleration in oblique electric potential structures (Greenspan and Whipple, 1982). Recently observations from within the ion conic source region (Kintner and Gorney, 1983) indicate that perpendicular electric field amplitudes in each of these models may be somewhat smaller than those required by present theories to account for the observed particle acceleration.

### Magnetotail

Plasma populations in the Earth's magnetotail show significant spatial and temporal variability in which the general pattern is only now becoming clear. The major regions of the Earth's magnetotail are (Eastman et al., 1983a, b): lobe, plasma sheet boundary layer and central plasma sheets, as illustrated in Figure 2-4.

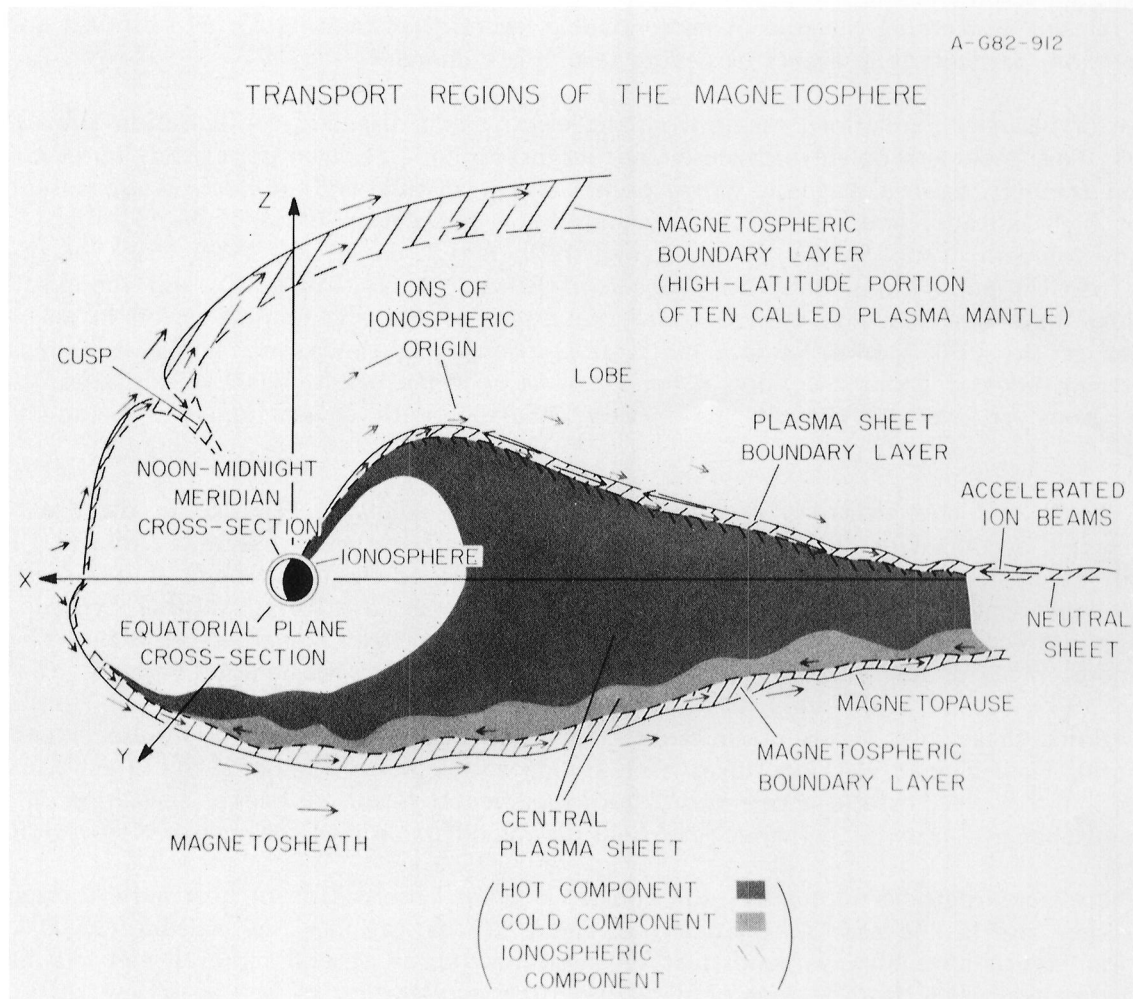


Figure 2-4. Particle acceleration geometry in the terrestrial magnetotail (courtesy Eastman et al., 1983).



The lobe has the lowest population with number densities ranging from less than  $0.01 \text{ cm}^{-3}$  to a few tenths  $\text{cm}^{-3}$ . The lobe plasma, at least in the neighborhood of the plasma sheet appears to consist of particles of ionospheric origin with  $\text{O}^+$  constituting up to one-half of the population. The ions are generally streaming away from the Sun with energies of approximately 100 eV. Temperatures of ions and electrons are very low. The lobe populations are also seen to interpenetrate the plasma sheet population.

The plasma sheet boundary region is about 1000 km thick and is characterized by earthward streaming ions with energies of a few to tens of keV, with streaming velocities of 100's of km/sec. Densities are in the range of  $0.1$  to  $1.0 \text{ cm}^{-3}$ . Beams away from the Sun are seen, but they are believed to be particles that have magnetically mirrored. The plasma sheet boundary layer is also a region of intense, broad-band electrostatic noise. It is tempting to attribute this noise to the interaction between the streaming and counter-streaming beams.

As one moves inward toward the central plasma sheet, counter-streaming ion beams are observed, which give way to a largely thermalized population characterized by ion temperatures on the 0.5 to 5 keV range and with streaming speeds generally less than 100 km/sec. In addition to the ions, the central plasma sheet is also populated by 0.1 to 1.0 keV electrons. In a related observation, Gurnett et al. (1976) report the observation of intense, nearly perpendicular electrostatic waves in the frequency range between the lower hybrid resonance and the electron gyrofrequency and in regions of large magnetic field gradients. It is tempting to attribute the heating of electrons to these waves.

There seems to be some disagreement on the relationship between temporal variation in plasma sheet populations and auroral substorms. Akasofu (1977) has identified the plasma sheet thinning with substorm onset, which is later followed by a late plasma sheet expansion and streaming toward Earth during the recovery phase of the substorm. Eastman et al. (1983a, b) maintain there is no clear relation between plasma sheet morphology and the occurrence of substorms, except that the percentage of  $\text{O}^+$  is enhanced during periods of high magnetic activity.

The ring current and radiation belt populations in the inner magnetosphere appear to be the result of inward convection of these central plasma sheet particles. There is yet another particle population in the Earth's magnetotail, namely the very energetic electron and proton bursts in which protons of energies in excess of 1 MeV and electrons of several hundred keV are observed (Krimigis and Sarris, 1980; Belian et al., 1981). In spite of the very high energy of individual particles, the total energy content of these bursts is small in comparison to the particle populations described above. More energetic proton bursts are seen on the evening side of the magnetotail while more energetic electrons are observed on the dawn side. However, since the particle energies are large in comparison to the potential drop across the magnetotail (typically  $\sim 50$  keV), energization of these particles cannot be simply attributed to the particles falling through a potential drop. The inferred acceleration regions appear to be localized, and there could be multiple acceleration regions (cf., Baker et al., 1979). Acceleration time scales range from  $<10$  seconds to several minutes. Although a number of suggestions have been made (cf., Galeev, 1982), there does not appear to be any generally accepted model of these bursts.

### **Bowshock**

The observation of low energy (a few keV, Asbridge et al., 1968), medium energy ( $\sim 30$  to  $\sim 100$  keV, Lin et al., 1974) and more energetic (up to  $\sim 1$  MeV, Sarris et al., 1976) ions upstream from the Earth's bowshock suggests this region of space (generally referred to as "foreshock") as a probable site for investigation of possible acceleration processes. These early particle observations have expanded enormously by the comprehensive measurements of ISEE-1, -2, -3 and have been supplemented by comprehensive wave measurements. In addition, similar phenomena have now been identified and studied at Jupiter (Zwickl et al., 1980) and Saturn (Krimigis et al., 1983).



Observations of these phenomena at Jupiter's magnetosphere have shown gross similarities to those at Earth (Zwickl et al., 1981), and have provided some important diagnostics on the source. It appears that a good part of the ion population at Jupiter consists of oxygen and sulfur ions, and it is also relatively poor in helium. Since oxygen and sulfur are overabundant in Jupiter's magnetosphere (Hamilton et al., 1981), it follows that this component of the upstream population is clearly of magnetospheric origin. Thus, the similarities and differences between Earth and Jovian observations may be instrumental in identifying possible acceleration processes operating in situ from those occurring within the magnetospheres of the planets.

To recapitulate, from the particle acceleration point of view, the foreshock is particularly relevant because of its plethora of structure in particles, waves and fields. The basic question that theory must address is the degree of in situ acceleration, as compared to leakage of already accelerated energetic ions from the nearby reservoir of the magnetosphere. This question may have more than one response, in that the detailed answer may well depend on which part of the energy spectrum one addresses, the state of the upstream solar wind and IMF parameters, and the state of the magnetosphere itself. Some of the issues will be discussed in Section III.

## PARTICLE ACCELERATION OUTSIDE THE SOLAR SYSTEM

Studies of particle acceleration in solar terrestrial environments are related to problems of particle acceleration in astrophysical systems in several important ways. The wealth of information available about energetic particles and sites of acceleration in the interplanetary medium and terrestrial magnetosphere makes it possible to develop well-constrained, testable theories of particle acceleration in these systems. The resultant insights into acceleration processes can then be profitably applied to problems of particle acceleration outside the solar system. Such exercises in extrapolation are not without benefit to the solar terrestrial community. The very high particle energies encountered in astrophysics, and the wide range of conditions under which acceleration apparently occurs provide a challenging test of our understanding of how acceleration mechanisms scale with the properties of the source, and the range of validity of acceleration theories.

Galactic cosmic rays (GCR) are the most thoroughly studied extrasolar system energetic particle population. A recent review of GCR, with orientation to the acceleration problem, was given by Axford (1981). GCR are observed by direct detection of the particles at Earth, by the synchrotron radiation of relativistic electrons in the galactic magnetic field, and by the  $\gamma$ -rays resulting from collisions between cosmic rays and ambient interstellar particles.

The GCR energy spectrum extends from at least  $10^{19}$  eV down to the MeV range, and can be represented by a power law with  $n(E) \sim E^{-2.6}$  particles  $\text{cm}^{-3} \text{GeV}^{-1}$  for  $10 \text{ GeV} < E < 10^6 \text{ GeV}$ . The spectrum steepens for  $E \gtrsim 10^6 \text{ GeV}$  and possibly flattens for  $E > 10^{10} \text{ GeV}$  (Cesarsky, 1980). The GCR energy density in the interstellar medium is about  $1 \text{ eV cm}^{-3}$ , comparable to the thermal gas and magnetic field energy densities. Protons outnumber electrons at the same energy by about 100 to 1. The residence time of GCR in the galaxy is estimated from spallation products to be about  $10^7$  year, with an energy dependence proportional to  $E^{-0.4}$ . The latter result implies that the energy spectrum at the source is flatter than the observed spectrum,  $f(E) \sim E^{-2.2}$ . Except for an overabundance of light elements, which can be explained by spallation of heavier nuclei, the elemental composition of GCR does not greatly differ from that of the SEP (cf., Section on "Solar Flares"); there are, however, important differences in isotopic composition (e.g., Ne; cf., Garcia-Munoz et al., 1979, and Mewaldt et al., 1980).

The origin of GCR is still unknown, but the most popular current models involve the acceleration of particles at shocks. The shocks might be produced by expanding supernova remnants (Blandford and Ostriker, 1980) or by stellar winds (Casse and Paul, 1980). Acceleration by a

steady, quasi-parallel, planar shock produces approximately the correct energy spectrum, and the resulting GCR composition would presumably resemble that of the interstellar medium; however, there is no generally accepted explanation of the observed isotopic anomalies. Some important points which must still be elucidated are: (i) the injection mechanism -- is a superthermal seed population required? (ii) the mediation of the shock structure by the accelerated particles themselves, (iii) the effect of curvature and temporal evolution of the shock on particle acceleration, (iv) the processes which determine the particle diffusion coefficient in the shock vicinity, and (v) the behavior of electrons -- can they be produced by shock acceleration as well?

### III. THEORETICAL PARTICLE ACCELERATION MECHANISMS

There is a rather substantial review literature on the subject of theoretical models for particle acceleration (cf., Arons et al., 1979; Axford, 1981; Forman et al., 1984; and references therein). The following discussion summarizes the principal results reported to date, with an emphasis on acceleration processes germane to the solar-terrestrial context.

#### ADIABATIC COMPRESSION, MAGNETIC PUMPING, AND DIFFUSION

Particles gain energy in gyrational motion through betatron acceleration when the magnetic flux through the orbit increases. The particle energy increases in a way that maintains the constancy of the magnetic moment

$$\mu = W_{\perp} / B \quad (2)$$

where  $W_{\perp}$  is the perpendicular kinetic energy and  $B$  is the magnetic field intensity ( $\mu$  is the (first) adiabatic invariant if the time scale of magnetic field changes is large when compared to the particle gyroperiod). Of course, the total particle energy increases only if an electric field is present; a static, inhomogeneous  $\vec{B}$  field does not work on particles (so that the partitioning of the kinetic energy of a given particle between the parallel and perpendicular components is governed simply by Equation (2)).

The parallel kinetic energy of trapped particles is increased when the particle mirror points come close together; this process is often referred to as Fermi acceleration. In this case, the parallel energy increases such that the parallel (second) adiabatic invariant

$$J = \int p_{\parallel} d\ell \quad (3)$$

remains constant, provided the time scales are long in comparison to a bounce period.

Both processes may be combined under a single heading, adiabatic compression, which is the result of a decrease in the flux tube volume,

$$V_B = \int d\ell / B \quad (4)$$

where the integral is along a field line path between particle reflection points. The effect is analogous to the compressional heating of a monatomic gas, with the parallel (perpendicular) temperature increasing as for a gas with one (two) degree(s) of freedom.

The best-known example of adiabatic compression of plasma is when an electric field convects plasma in a planetary magnetosphere from a low to a high magnetic field region; here, most of the kinetic energy gain occurs in the perpendicular component, and is equal to the loss of electric potential in the convection electric field. Such a process is responsible for energizing the radiation belt electrons and ring current protons in the Earth's magnetosphere. The effects are seen in the generation of whistler modes ("chorus"; cf., review by Thorne, 1972) and in electron precipitation, which causes a diffuse auroral glow, particularly in the morning sector of the auroral oval.

There are three factors that modify this process. One is corotation of the magnetic field. Corotation may be considered to be an azimuthal  $\vec{E} \times \vec{B}$  drift which, in the equatorial plane, implies a radially-directed electrostatic field; in the case of the Earth, the corotational potential is negative, so that ions gain energy from the corotational potential. Another factor influencing inward convection of radiation belt particles is the polarization of a plasma cloud that is being convected inward; this is due to the opposite drifts of ions and electrons. If there is nothing to discharge the polarization electric field, it will completely cancel the imposed convection electric field, shutting off convection into regions of smaller flux tube volume. As has been shown by Harel et al. (1981) in the case of the Earth's magnetosphere, the presence of a conducting ionosphere at the base of field lines will partially discharge this polarization electric field to allow the inward convection of plasma. Inward convection of ring current plasma is strongly dependent on auroral ionization, which greatly enhances the conductivity of the nightside ionosphere. Enhanced auroral activity allows the deeper penetration of ring current protons. The final factor affecting acceleration through adiabatic compression is the time variation of the convection electric field or of the trapping magnetic field on time scales long in comparison to a particle bounce time, but short in comparison to the time it takes a particle to drift around the Earth; for a review of this process see Walt (1971). Basically, an azimuthally asymmetric perturbation of a dipole field causes inward and outward displacement of particle orbits. If the perturbation undergoes a change in the time it takes a particle to make a complete circuit around the dipole source of the trapping field, a net displacement of the particle orbits will have taken place. Repeated perturbation has the effect of diffusing the particle across magnetic L-shells, i.e., breaking of the third adiabatic invariant. The diffusion coefficient for magnetic perturbation has a  $L^{10}$  dependence, while electric field fluctuations produce a  $L^6$  dependence, where L is the geocentric distance in Earth radii where the particle guiding center intersects the equatorial plane. Even though electric field fluctuations for a given power spectrum may be more effective in causing diffusion at low L values, electric field fluctuations inside the ring current are to a large extent shielded by the ring current. Lanzerotti and Wolfe (1980) have concluded that in the inner magnetosphere, radial diffusion caused by electric fluctuations is comparable to that caused by magnetic fluctuations. These fluctuations allow a small portion of the ring current particles to diffuse to geocentric distances considerably inside the ring current belt, and thereby to obtain much higher energies.

A somewhat distinct process of particle acceleration involving the adiabatic invariants is magnetic pumping. This process combines a cyclic variation of the magnetic field with isotropization of the particle distribution during intervals of maximal field strength (Alfvén and Falthammar, 1963). As the magnetic field strength increases, the perpendicular momentum increases as well (since the first adiabatic invariant, Equation 2, is conserved), but the parallel momentum is conserved. If isotropization of the distribution function occurs systematically when the field strength is near maximum, then the perpendicular momentum will be systematically increased (since it is conserved whenever isotropization does not occur); hence, the total kinetic energy of particles will increase secularly.

The basic physical issue involved in magnetic pumping is the mechanism of isotropization. Several candidate processes have been invoked in the literature, including short wavelength MHD modes which scatter efficiently and whistlers (Thorne, 1983; Kennel and Ashour-Abdalla, 1982; Kulsrud, 1979). The origin of these scattering waves itself is subject to uncertainty: these waves may owe their existence to the accelerated particles themselves, or may be generated by processes incidental to the acceleration mechanism itself. The interested reader may wish to consult Goertz (1978) for an example of an application of the magnetic pumping mechanism (in the context of particle acceleration in Jupiter's magnetosphere).

## ACCELERATION IN DIRECT ELECTRIC FIELDS

A particularly well-studied mechanism for particle acceleration is by direct electric fields. Such fields are associated with magnetic reconnection in the vicinity of magnetic neutral points and current sheets (Syrovatskii, 1981), where they appear perpendicular to the magnetic field. Particle acceleration is also possible in electric fields parallel to the magnetic field (Colgate, 1978), which can arise from the interruption (due to plasma instabilities) of the parallel currents associated with twisted magnetic flux tubes and from the formation of double layers of electric charge (e.g., Block, 1972; Carlqvist, 1979; see Section on "Double Layers" below).

The application of direct electric fields to the acceleration of non-relativistic electrons in solar flares has been discussed, among others, by Smith (1979) and Heyvaerts (1981). Here we wish to emphasize the role that direct electric fields could play in the very rapid production of nonthermal ions and electrons in solar flares. As discussed in the Sections on "Stochastic Acceleration" and "Shock Particle Acceleration", both stochastic and shock acceleration could accelerate ions quickly enough to account for the gamma-ray observations, but it is not clear that the mean free paths are, in fact, sufficiently short and the turbulent and shock velocities are large enough to account for the rapid acceleration. On the other hand, in an electric field model (such as that of Colgate, 1978), the acceleration time could be as short as 0.1 second for a loop of length  $\sim 10^9$  cm. Relativistic electrons around 10 MeV, however, may not be produced fast enough and in sufficient quantities to stochastic and shock acceleration because the particle energy gain per collision is proportional to particle mass. But there seems to be a good correlation between interplanetary electrons of such energies and gamma-ray flares (Evenson et al., 1981; see Section II). Since the gamma rays are produced by ions of  $\sim 10$  MeV/nucleon, this correlation could indicate electron and ion acceleration to the same energy by an electric field.

In this section, we review particle acceleration processes in current sheets, in the vicinity of magnetic neutral points, and by parallel electric fields.

## Current and Neutral Sheet Particle Acceleration

Current sheets and magnetic neutral points are involved in a wide variety of contexts in astrophysics, but for brevity's sake, the following discussion focuses primarily on results obtained in the solar-terrestrial context; more detailed discussions of neutral sheets and reconnection processes may be found in this volume in the Reconnection Working Group chapter (see also Syrovatskii, 1981 and Priest, 1983). Some of the present material is adapted from Forman et al. (1984).

The astrophysical interest in current and neutral sheets arises because of the variety of situations in which magnetic reconnection is thought to play a central role in the overall energetics; indeed, Parker and others (see Parker, 1979 and references therein) have emphasized the very special conditions required for magnetostatic equilibrium, and have proposed that reconnection of tangled magnetic fields occurs routinely and inevitably in nature. For example, it is believed that the primary energy release in solar flares occurs by magnetic reconnection (e.g., Kahler et al., 1980; Sturrock, 1980); hence, it is natural to consider particle acceleration mechanisms

associated with magnetic field dissipation and reconnection (see, for example, the discussion of reconnection and associated particle acceleration in the Earth's magnetosphere by Stern and Ness, 1982).

*General Considerations:*  $\vec{B}_\perp = 0$ .

In models of steady-state reconnection, magnetic fields of opposite polarity are brought together by a flow of fluid. A detailed discussion of the fluid theory of steady-state reconnection has been given by Vasyliunas (1975); two of the best known examples of reconnection are the models of Petschek (1964) and Sonnerup (1970) which involve x-type neutral points. These field lines lie in a plane (say, the x-y plane) and there is a constant electric field,  $\vec{E}$ , perpendicular to that plane. The electric field is related to the magnetic field  $\vec{B}$ , flow velocity  $\vec{V}$ , and current  $\vec{j}$  in the plasma by Ohm's law.

$$\vec{E} + \vec{V} \times \vec{B} / c = \eta \vec{j}, \quad (5)$$

where  $\eta$  is the resistivity. The inclusion of other terms in Equation (3.4) is discussed by Vasyliunas (1975). The resistive term is unimportant everywhere except in a small region near the x-type neutral point where  $\vec{B}$  becomes small and  $\vec{j}$  is large (because the gradients of  $\vec{B}$  are large). Here the frozen-in condition of the magnetic field is violated. Rather than being transported by the fluid, the field diffuses through the semi-stagnant plasma. The region, therefore, is referred to as the diffusion region.

In this model, cold fluid flows inward along the x-axis. Most of the fluid never enters the diffusion region, but passes through a shock front (or pair of fronts, in Sonnerup's model) and flows out parallel to the y-axis. In Petschek's model, the inflow occurs at a small fraction ( $\sim 0.1$ ) of the Alfvén speed, while the outflow is exactly Alfvénic. This acceleration takes place at the expense of magnetic energy. In Sonnerup's model, the inflow can be faster than in Petschek's model, in which case less acceleration of the bulk flow occurs. In both models, however, the thermal speed of the ions increases across the shock front to about the Alfvén speed. The increase of electron thermal energy depends on the structure of the shock and the nature of the coupling between electrons and ions.

Similar results have been obtained by Hayashi and Sato (1978), who have simulated time dependent, compressible reconnection by solving the fluid equations numerically. In their model, an initially planar neutral sheet is compressed into an x-type neutral line by imposed fluid flow directed inward, perpendicular to the initial field lines. They assume that when the current density  $j$  exceeds some critical value  $j_c$ , anomalous resistivity develops (see the Section on "Anomalous Resistivity" below for details), and increases as a power of  $j - j_c$ . Thus, reconnection of the magnetic field lines and Joule heating of the plasma near the neutral line begin when the neutral sheet has been sufficiently compressed by the incoming flow. They find that, as in the analytic models of reconnection, the regions of inflow and outflow are separated by shocks and that the outflow speed is on the order of the Alfvén speed.

If ions gain flow velocities on the order of the Alfvén velocity, their energy per nucleon is  $2.5 \times 10^4 B^2 / n$ , where  $B$  is in gauss and  $n$  in  $\text{cm}^{-3}$ . While for most combinations of  $B$  and  $n$  in the solar atmosphere, this energy is not as high as the observed particle energies, this bulk flow acceleration could nonetheless be important as an injection mechanism (whether observations demand such a "two-step" process is another matter, and is discussed in the Section on "Solar Flares" above). In certain exceptional cases, however,  $B$  could be high enough and  $n$  low enough for the particles to achieve energies as high as 10 MeV/nucleon (e.g., Sonnerup, 1973), which would suffice, for example, for the production of nuclear  $\gamma$ -rays.

In addition to the acceleration and heating of the plasma as it passes through the reconnection region, the possibility exists of direct particle acceleration in the electric field at the magnetic null line. An example of such acceleration was discussed by Speiser (1965). He considered a magnetic field with field lines parallel to the  $y$  axis, with field reversal at  $x=0$ ,

$$\vec{B} = \hat{e}_y B_1 - \frac{x}{d} \hat{e}_z \quad |x| < d \quad (6)$$

$$\vec{B} = \hat{e}_y B_1 \quad |x| > d \quad (7)$$

where  $\hat{e}_y$  is the unit vector in the  $y$  direction and  $B_1$  is a constant. Hence the field equals  $B_1$  for  $|x| > d$ , and goes linearly to zero for  $|x| < d$ . A constant electric field  $\vec{E} = VB_1 \hat{e}_z / c$  is also present, and is assumed to be continuous at  $x=0$ . Everywhere in space except near the field reversal region at  $x=0$ , the particle orbits consist of gyration in the  $x$ - $z$  plane (the usual Larmor motion), combined with the  $\vec{E} \times \vec{B}$  drift in the  $x$  direction toward the plane  $x=0$ . Since for  $|x| > d$  the motion in  $z$  is oscillatory, the average energy gain is zero in this region. However, once a particle has drifted to near the  $x=0$  plane where the field is very weak, it does not gyrate - it is simply accelerated in the  $z$  direction. The final energy,  $E$ , of the particle depends on the extent,  $L$ , of the current sheet in the  $z$  direction. Numerically,  $E$  equals  $10^{-6} BLV$  MeV, where  $B$  is in gauss,  $L$  in km and  $V$  in km/sec. Thus, particles can be accelerated to tens of MeV if  $B$ ,  $V$  and  $L$  are large enough (with the energy gain of course limited by the total potential drop across the neutral sheet). Whether neutral sheets have sufficient lengths, and whether or not particles can remain in the sheet for such lengths, are problems which have not yet received clear-cut answers, at least for solar flares.

In order to estimate the distance the particles stay in the neutral sheet, particle orbits in various magnetic geometries have been calculated by Speiser (1965), Friedman (1969), Bulanov and Sasorov (1976), Bulanov (1980), and Syrovatskii (1981). The results depend strongly on assumed geometries, and analytic treatments can only be made for the simplest field configurations. Time dependent electric and magnetic fields have also been considered (Burke and Layzer, 1969; Levine, 1974). Levine extended Burke and Layzer's work by trying to estimate the effect of Coulomb collisions between the test particles and field particles in the ambient medium. The existence of an energy loss mechanism which competes with acceleration sets an upper limit on the time constant for collapse of the neutral sheet such that particle acceleration can still occur. Recently, Mullan and Levine (1981) have studied the implications of the collapsing magnetic neutral sheet model for the composition of solar flare accelerated ions.

Attempts have also been made to calculate the energy spectrum of particles accelerated in neutral sheets. In the absence of stochastic processes such as Coulomb collisions or wave-particle interactions, the phase space distribution function  $f$  of the test particles satisfies Liouville's theorem; the density of particles in phase space is constant on phase space trajectories. Therefore, given an initial spatial distribution of injected particles and solutions for the particle orbits, the momentum distribution function of particles leaving the acceleration region can be calculated (Bulanov and Sasorov, 1976; Bulanov, 1980). They find that when the initial distribution of particles is uniform in space, the spectrum is an exponential in energy or in a fractional power of the energy. Energy spectra have also been calculated numerically by Bulanov (1980), who finds good agreement with the analytic results. Friedman (1969) has used numerical techniques to calculate final energies for test particles in Petschek's (1964) reconnection model. A generally recognized problem with particle acceleration in neutral sheets is that only very few particles are accelerated, while the injection of particles into the acceleration region (see Section on "Coherent Wave-Particle Acceleration") is specified ad hoc in the models.



*A Particular Example:  $\vec{B}_\perp \neq 0$ .*

If the magnetic field component perpendicular to the magnetic field reversal plane does not vanish, particle acceleration can also occur, but to energies somewhat below those achievable by the strictly neutral sheet model discussed above. A mechanism of this type with particular magnetospheric relevance is ion acceleration by perpendicular electric fields in regions where magnetic field lines have large curvature, and when there is an electric field perpendicular to the plane containing the magnetic field lines (see Figure 2-6).

To understand this mechanism, we can transform to an inertial frame translating to the left at a speed of  $u_x = c E_y / B_z$ , where  $E_y$  is the uniform electric field directed out of the page of Figure 2-6. In this frame, a low-energy ion will then be streaming to the right along the field lines at a speed of approximately  $u_x$ . When it encounters the region of high curvature, its field-aligned motion will be reversed upon ejection, either above or below the region of high curvature. Back in the observer's frame, the velocity of translation  $u_x$  will be added to the particle's motion, so that a particle will have gained an energy in the range of  $\mu u_x^2$  to  $\sim 2\mu u_x^2$ . Taking conditions typical of the Earth's magnetosphere of  $E_y \approx 1$  mV/m and  $B_z = 5$  V, the energy gain,  $\mu u_x^2$ , is 400 eV for  $H^+$  and 6 keV for  $O^+$ .

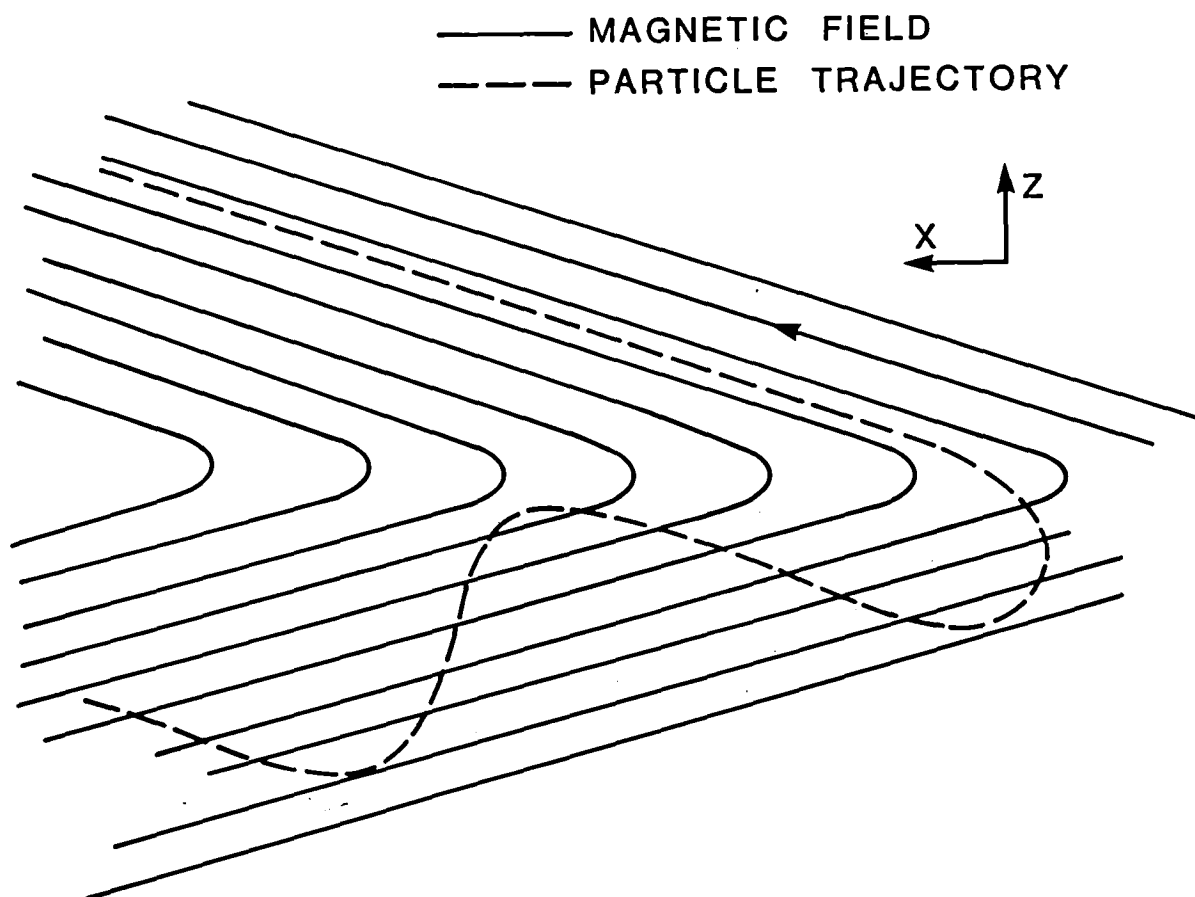


Figure 2-6. Geometry appropriate for ion acceleration by electric fields perpendicular to the plane containing magnetic field lines (Section entitled "Acceleration in Direct Electric Fields").

Swift (1977) and Wagner et al. (1979) have studied individual particle trajectories to demonstrate this mechanism as a process for energizing ions in the region of the neutral sheet of the magnetosphere, assuming the presence of a convection electric field. Subsequently, Lyons and Speiser (1982) have integrated the orbits of a large number of particles to demonstrate that the process can account in detail for the velocity distribution of ions observed in the streaming plasma that forms the boundary of the plasma sheet. Swift (1982, 1983) has also shown that, by assuming that the lobes of the Earth's magnetotail are populated by low energy ions of ionospheric or magnetosheath origin, the momentum transferred to these particles upon encountering the neutral sheet can be sufficient to balance the Maxwell stresses associated with the highly extended magnetotail of the Earth's magnetosphere. Hence, this acceleration process appears capable of accounting for energizing the ions in the Earth's plasma sheet and maintaining the magnetosphere's extended magnetotail. The acceleration process does not depend on whether the electric field is of external origin or due to an induction field generated by the relaxation of the magnetic field. Hence this acceleration process can occur when the system is externally driven or when the energy source is the magnetic field energy stored in the system.

### Particle Acceleration by Parallel Electric Fields

When there exists a component of an electric field parallel to a magnetic field, the acceleration processes are obvious: ions are accelerated along the magnetic field in the direction of decreasing potential, while electrons are accelerated in the opposite direction; if there is no interaction of the accelerated particles with ambient particles or interaction among the accelerated particles, the energy gain is equal to the potential through which the particles have moved. Two major questions concerning acceleration by parallel electric fields are: (1) what process generates the potential; and (2) what limits the current so the potential will not be shorted out? We shall not discuss the problem of generating field-aligned potentials, but will rather focus in the following on mechanisms for current limitation that allow a (given) potential to be maintained.

#### *Ion Inertia.*

Currents can be limited by a variety of processes. An obvious process is a lack of charge carriers and ion inertia. This process appears to be operative over the terrestrial polar cap, where downward pointing electric fields were observed by Heppner et al. (1981). Because the polar cap field lines link into the lobes of the magnetotail where the plasma densities are very low, there are few electrons available to short out the potential; a major portion of charge carriers is therefore likely to be ions. At present there is no quantitative model that links the upward flux of ions and the energy gain of ions with the convection electric field.

#### *Anomalous Resistivity.*

The term "anomalous resistivity" refers to a class of microprocesses which increase the effective scattering rate of current-carrying charges in a plasma (and hence increase the thermalization rate of free energy carried by the current flow). The basic idea underlying the occurrence of anomalous resistivity is that a current-carrying plasma can be thought of as two interpenetrating beams; if the combined electron and ion distribution function is "double-humped", it may then be unstable to the generation of plasma wave modes, which in turn act to impede the flow of ions and electrons, and hence limit the current that can be carried (cf., Papadopoulos, 1977).

The most easily excited instability in a current-carrying plasma is the electrostatic ion cyclotron wave or, if the electron temperature is much greater than the ion temperature, the ion acoustic wave. The critical drift speed for onset is approximately the ion thermal speed (Kindel and Kennel, 1971). With these instabilities, it takes only an energetically very small change in the

electron distribution function to stabilize the velocity distribution; hence, these instabilities may not be effective in inhibiting the current flow. It is therefore very useful to be able to estimate the anomalous transport coefficients; these can, for example, be simply estimated by assuming marginal stability for the combined distribution function (cf., Manheimer and Boris, 1977).

Anomalous resistivity is likely to be significant only when the electron flow rate becomes comparable to the electron thermal speed to excite the electron plasma in Buneman modes. The  $\mathbf{J} \cdot \mathbf{E}$  power dissipated will go into heating the electrons, which raises the instability threshold. Thus, the effect of anomalous resistivity is to allow the partitioning of electron energy among streaming and internal degrees of freedom, but is not likely to result in much more than about a factor of two reduction in electron mobility over electron inertia.

An interesting variation on the anomalous resistivity scheme (Papadopoulos and Coffey, 1974) has been proposed by Rowland et al. (1981). It is assumed that a large amplitude, low-frequency wave traps low-temperature electrons. Some electrons escape trapping and are accelerated by the electric field; these electrons interact with the stationary electrons and excite electrostatic waves associated with the upper hybrid resonance. These waves are capable of distributing streaming energy of the runaways among all three internal degrees of freedom, and can substantially impede the runaways (the "runaway tail" problem occurs whether or not one has Coulomb or anomalous resistivity, and results from the strong inverse power law dependence of the electron collision cross-section on velocity for high energy particles; see Heyvaerts, 1981 for a discussion in the solar flare context). In order that the majority of electrons be trapped, ion waves must satisfy  $\delta n_i/n > 0.2$ . In the aurora, ion-cyclotron waves have been measured to have large amplitudes (Mozer et al., 1977; Kintner et al., 1978); these waves have also been observed with large amplitudes ( $\delta n_i/n > 0.5$ ) in laboratory experiments (Bohmer and Fornaca, 1979) and computer simulations (Ashour-Abdalla et al., 1981), and are driven by a low velocity drift between electrons and ions ( $V_d < V_{te}$ ). The recent laboratory experiments of Lang and Bohmer (1983) have clearly shown that ion-cyclotron waves generated by a low velocity electron current ( $V_D < V_{te}$ ) can trap a significant fraction of the electrons (>50%), in good agreement with the model; Lang and Bohmer also observed a parallel d.c. electric field when the ion-cyclotron turbulence was present (see Section on "Auroral Particle Acceleration" above and "Double Layers" below). Thus, the anomalous resistivity mechanism may after all significantly inhibit parallel current flow in the magnetospheric context, but this remains to be established in detail.

In any case, the above does not mean that anomalous resistivity may be unimportant. For example, in the Section on "Auroral Particle Acceleration" we discussed the presence of energetic ions above the Earth's ionosphere; these ions can be heated by ion cyclotron waves (Dusenbury and Lyons, 1981). In order that there be a continuing transfer of energy to the ions, there must be some anomalous resistivity to maintain a small electric field so that  $\mathbf{J} \cdot \mathbf{E}$  is nonzero. A definitive investigation of the effectiveness of this process by numerical simulation has yet to be carried out.

Finally, it should be pointed out that in any particle acceleration problem involving parallel electric fields and currents, it is crucial to understand the global structure of the physical system. This point is very simply illustrated by noting that if the electric field  $E_{\parallel}$  is fixed, the heating rate  $\mathbf{J} \cdot \mathbf{E} = \sigma E_{\parallel}^2$  is directly proportional to the conductivity, while if the current density  $J_{\parallel}$  is fixed, the heating rate  $J_{\parallel}^2 / \sigma$  is inversely proportional to the conductivity. Evidently, it is of considerable consequence to determine whether the global system behavior fixes  $E_{\parallel}$  or  $J_{\parallel}$ .

### *Double Layers.*

Another variation on the anomalous resistivity theme is that a plasma which becomes unstable due to current flow develops a large amplitude voltage jump called a double layer (Block, 1972; Carlqvist, 1979). This had been a popular, but controversial, concept for over a decade. Analytical work and numerical simulations suggest that individual double layers are likely to be impermanent features; for example, one-dimensional numerical simulations in "long" ( $4,096 \lambda_D$ ) systems for both periodic and nonperiodic boundary conditions carried out by Sato and Okuda (1981) show that a current-carrying plasma will support multiple, small double layers having a transitory existence (see their Figures 3 and 10); these simulations indicate ion-acoustic double layers on the order of  $50 \lambda_D$  thickness (with  $e\delta\sigma/T_e \approx 1$ ) spaced about  $1,000 \lambda_D$  apart. More recently, Temerin et al. (1982) have reported the observation of a small transitory double layers about 6,000 km above the Earth's ionosphere with a maximum amplitude of about 10 mV/m. The fluctuation amplitude required so that the majority of electrons is trapped is  $\delta n_i/n > 0.2$ ; if we recall the discussion of the previous Section immediately above, it would seem that electron trapping (in that case, by ion-cyclotron waves generated by relatively low electron drift speeds,  $V_D < V_{te}$ ) could be a significant source of effective resistivity in the magnetosphere (see, however, Swift, 1981, for a more cautionary treatment).

It remains to ask whether the total potential drop across multiple double layers can account for a significant portion of the overall parallel potential drop. Adopting the electric field value given by Temerin et al. (1982), and assuming it is distributed over 1/20 of the vertical height range of 4,000 km, a maximum potential drop of 2 kV could be obtained; this value is consistent with auroral precipitation energies. However, although the microdouble layer model may in principle account for a substantial potential drop (and, indeed, microdouble layers have been observed), there have also been instances when particle data have indicated the presence of a parallel electric field in the region of the observing satellite, but no double layers have been observed. It is therefore likely that an additional process for maintaining a field-aligned potential drop is needed.

### *Magnetic Mirror Forces.*

Potential differences can also be maintained by magnetic mirror forces on energetic particles. In the case of the Earth's magnetosphere, currents driven by an upward electric field are inhibited by inertia of ionospheric ions and upward directed magnetic mirror forces on energetic electrons. The basic ideas were developed by Lemaire and Scherer (1971) and Knight (1973), and later calculations by Lyons (1980) and Chiu and Cornwall (1980) were used to account for specific auroral features.

One limitation of the magnetic mirror model is that the maximum potential drop is limited to approximately the thermal energy of the electrons above the region of field-aligned electric field. However, simulations carried out by Wagner et al. (1981) have indicated that potential drops large in comparison to the electron thermal energy can be maintained in a converging field. The new element added by the simulations was the presence of trapped electrons between the potential and magnetic mirror. In the construction of an analytic model, the trapped electrons permit the maintenance of quasi charge neutrality in the region below a large potential drop. Since these trapped populations are observed in regions of auroral precipitation, it appears that the critical condition for this mechanism is fulfilled. Moreover, simulations indicate that this process is not subject to disruptive instabilities. It now must be shown that such a model yields electron spectra consistent with that observed above auroral forms.

### *Perpendicular Structure.*

One of the most obvious features of the aurora as seen on Earth is the fine-scale structure perpendicular to magnetic field lines. The question naturally arises as to whether there is any feature in the current-limiting mechanism that can account for this structure. Although the micro-double layer and magnetic mirror mechanisms can accommodate perpendicular structure (Chiu and Cornwall, 1980), there is nothing in them that demands it. Swift (1979) has formulated a model for static V-shaped potentials, but was unable to find a self-consistent solution to the coupled Poisson-Vlasov system for a V-shaped potential (of the type shown in Figure 2-3). However, the theory indicates that there is a minimum perpendicular size and hence a minimum width of auroral forms. This minimum width turns out to be the gyrodiameter of an ion whose energy is equal to the potential difference across the V-shaped potential.

V-shaped potentials have also been investigated by numerical simulation by Borovsky and Joyce (1983) using a particle code. Their models assumed perpendicular size scales less than the ion gyrodiameter, and in each of their runs the parallel potential was dropped within a few Debye lengths of the upper boundary. This indicates potentials with small size scales would close very near the region in which they are generated. Hence there seems to be no satisfactory potential model that requires the perpendicular, as well as parallel, structure. The fine-scale auroral structure is thus more likely a consequence of the generating mechanism than of the current limiting process.

## STOCHASTIC ACCELERATION

As discussed by Forman et al. (1984), from which the following has been adapted, processes in turbulent plasmas which cause particles to change their energy in a random way -- with many increases and decreases in energy -- lead to stochastic acceleration. In the original stochastic Fermi mechanism (Fermi, 1949), the process was reflection from randomly moving magnetized clouds. Stochastic acceleration can also result from resonant pitch-angle scattering from electromagnetic waves with wavelength of the order of the particle gyroradius. To accelerate particles, these waves must propagate both parallel and anti-parallel to the average magnetic field (Skilling, 1975). Other modes of stochastic acceleration, called magnetic pumping (cf., Section on "Adiabatic Compression, Magnetic Pumping and Diffusion") and transit-time damping, occur through interaction with magnetosonic waves whose wavelengths are much longer than the particle gyroradius (Kulsrud and Ferrari, 1971; Melrose, 1980; Achterberg, 1981); these modes require additional pitch-angle scattering to keep the particles isotropic. Langmuir (plasma) waves or other electrostatic waves with phase velocities of the order of the particle speed will also accelerate particles stochastically (Melrose, 1980).

When the random energy increments are small compared to the particle energy, stochastic acceleration results in a diffusive current in momentum space,  $S_p = -D_{pp} \partial f / \partial p$ , where  $p$  is the magnitude of the momentum,  $f(p)$  is the number of particles per unit volume in phase space and  $S_p$  is measured in  $\text{cm}^{-3} \text{momentum}^{-2} \text{sec}^{-1}$ . Thus, particles injected at some momentum  $p_0$  will diffuse in momentum to larger and smaller  $p$ . In terms of  $f$ , the differential particle intensity per unit energy per nucleon is given by  $dJ/dE = A p^2 f$ , where  $A$  is the nuclear mass number. Additional non-diffusive energy changes can be added to  $S_p$ ,

$$S_p = D_{pp} \frac{\partial f}{\partial p} + \frac{dp}{dt} f \quad (8)$$

where  $dp/dt$  represents convection in momentum space due to processes which change the energy of all particles (e.g., ionization or Coulomb losses). The aim of theory is to determine the form of the momentum diffusion coefficient  $D_{pp}$ , which determines the efficacy of any given stochastic acceleration process; in the following, we briefly sketch the considerations which enter into the calculation of  $D_{pp}$ .

It is evident that the diffusion coefficient  $D_{pp}$  must depend on the nature of the stochastic process. In order to illustrate this point, we shall begin by considering the simplest case, in which particle scattering is assumed to be independent of pitch angle. In this highly simplified case, we think of the scattering process as hard sphere scattering with mean free path  $\lambda$ , and  $D_{pp}$  can then be simply derived from the Boltzmann equation (Parker and Tidman, 1958). This yields

$$D_{pp} = p^2(\delta V)^2/3v\lambda, \quad (9)$$

where  $(\delta V)^2$  is the mean square velocity of the scatters and  $V$  the particle speed.

Suppose the pitch angle dependence is not suppressed; consider then, for example, stochastic acceleration due to resonant pitch-angle scattering from Alfvén waves. In this case, the momentum diffusion coefficient obtained from quasi-linear theory is (Skilling, 1975)

$$D_{pp}^A = \frac{2p^2V_A^2}{v^2} \int \frac{D_+D_-}{D_++D_-} d\mu \quad (10)$$

where  $V_A$  is the Alfvén speed,  $D_+(D_-)$  is the pitch-angle scattering coefficient due to forward (backward) propagating Alfvén waves, and  $\mu$  is the cosine of the particle pitch angle in the mean field. It is clear from this equation that there is no stochastic acceleration due to Alfvén waves unless the waves propagate in both direction. This requirement occurs because the electric fields of Alfvén waves in one direction can be Lorentz-transformed away, and so cannot accelerate particles. For example, Alfvén waves generated by the streaming of energetic particles (Wentzel, 1969) propagate only in the direction of the streaming, and hence do not accelerate the particles.

By comparing Equations (9) and (10), we can define an effective mean free path for stochastic acceleration by Alfvén waves,  $\lambda^A = p^2V_A^2/3vD_{pp}^A$ ; this  $\lambda^A$  is, in general, a function of particle rigidity, which is determined by the power spectrum of the Alfvén waves. In order to evaluate the effective mean free path, let  $W_+(k)$  be the energy density per wave number  $k$  in waves propagating in the + or - directions; and use results from quasi-linear theory (Hasselmann and Wibberenz, 1968; Jokipii, 1971; Luhmann, 1976) to evaluate the diffusion coefficients  $D_+(\cdot)$ . Then, for example, if  $D_+ = D_-$  and  $W(k) \propto k^{n-2}$ , we can evaluate the effective mean free path,

$$\lambda^A(R) = \frac{B^2 r_c^2}{96\pi^2 W(1/r_c)} (2-n)(4-n) \propto R^n, \quad (11)$$

where  $r_c$  is the gyroradius of the charged particle in the average magnetic field,  $B$ .

A similar calculation for stochastic acceleration due to long-wavelength magnetosonic waves (assuming adequate particle pitch-angle scattering) leads to an expression for the diffusion coefficient  $D_{pp}$  (again using quasi-linear theory) of the form

$$D_{pp}^M = \xi \frac{p^2 V_A^2}{3V} \frac{\langle \delta B^2 \rangle}{B^2} \langle k \rangle \quad (12)$$

(adapted from Achterberg, 1981), where  $\langle \delta B^2 \rangle$  is the mean square of the fluctuations in the field magnitude and  $\langle k \rangle$  is the mean wavenumber of the magnetosonic waves. The number  $\xi$  depends on the angular distribution of the waves; it is usually assumed that  $\xi \sim 1$ . Note that in this case, the effective mean free path  $\lambda^M$  (corresponding to Equation 11) is given by  $B^2 \langle k \rangle^{-1} / \langle \delta B^2 \rangle$ , and is independent of particle momentum or charge. Note further that in this case, it is not necessary for the magnetosonic waves to propagate in both directions to accelerate particles. However, as already mentioned, a certain level of pitch-angle scattering is required to isotropize the particles; this latter condition is  $D_+ + D_- > D_{pp}^M / p^2$  (Achterberg, 1981).

The relative importance of acceleration by magnetosonic and by Alfvén waves is measured by the ratio  $D_{pp}^M / D_{pp}^A$ . In order of magnitude, this ratio is

$$\frac{D_{pp}^M}{D_{pp}^A} \approx \langle k \rangle r_c W^M \left( \frac{1}{k_r W_+(k_r)} + \frac{1}{k_r W_-(k_r)} \right), \quad (13)$$

where  $k_r = r_c^{-1}$  is the resonant wave number and  $W^M$  is the total energy density in the long wavelength magnetosonic waves. Since by assumption  $\langle k \rangle r_c \ll 1$ , acceleration by Alfvén waves dominates when their energy density is comparable to that in magnetosonic waves, but only when there is applicable power in Alfvén waves propagating in both directions. For the sake of completeness we also quote the form of the momentum diffusion coefficient for isotropic Langmuir turbulence,  $D_{pp}^L$ : this result is given by Melrose (1980; Equation 8.13), and shows that in this case the effective mean free path  $\lambda^L \propto (A/Z)^2$  times a function of particle velocity.

In the context of solar-terrestrial physics, the most detailed application of stochastic acceleration has been carried out for solar flares (Ramaty, 1979; Ramaty et al., 1983a; Forman et al., 1984). Here it has been shown that the Bessel function spectra derived from a simple nonrelativistic, steady state, acceleration model with constant diffusion mean free path and escape time can account for both the observed interplanetary spectra of protons and  $\alpha$  particles, as well as for the particle spectra at the Sun deduced from high energy neutron and  $\gamma$ -ray line observations. Many problems, however, remain. In particular, it is not clear whether or not stochastic acceleration is sufficiently rapid to account for the very fast rise times of the observed  $\gamma$ -ray emissions (Section II above) and, in addition, whether the necessary turbulence itself is developed on sufficiently rapid time scales, and to sufficiently large amplitudes. Furthermore, the question of the approximate constancy of the energetic particle spectrum is not satisfactorily addressed in this stochastic acceleration model.

## SHOCK PARTICLE ACCELERATION

Of all acceleration mechanisms, the shock process has been best studied both observationally and theoretically. As pointed out in Section II, examples of shock acceleration are firmly established in the interplanetary medium and upstream of planetary bowshocks, and to a lesser extent, in solar particle acceleration at the Sun and in astrophysical objects. In general, the observations suggest that shock ion acceleration operates over a large energy range (a few keV upstream of the Earth's bowshock to  $\sim 10^4$  GeV at supernovae envelopes) and for several ion species (observations of H, He,  $Z \geq 3$  are well established). A few examples of electron shock accelerations have also been reported in the interplanetary medium (Section II). The in situ observations have provided a wealth of information, including particle spectra and anisotropies, plasma wave and magnetic field signatures, and shock propagation parameters. These observations provide the most stringent tests of theoretical models, and have led to increasingly sophisticated analytic and simulation work in this area. We give below a short description of the basic shock acceleration mechanisms identified so far, and then proceed to point out those plasma environments where each is thought to be operating.

As seen by an observer moving with a magnetosonic fast-mode shock wave, the plasma enters (leaves) the shock front at a speed greater (less) than the fast-mode magnetosonic wave speed, and the flow speed on either side of the shock exceeds the Alfvén speed. In contrast, as seen by an observer moving with a magnetosonic slow-mode shock wave, the plasma enters (leaves) the shock front at a speed greater (less) than the magnetosonic slow-mode wave speed, and the flow speed on either side of the shock is less than the Alfvén speed.

The incoming plasma is decelerated, compressed and heated in the shock front over a distance  $\sim$ several thermal ion gyroradii. The very large gradients in  $\vec{B}$  and the plasma bulk velocity that occur at the shock front, combined with the induced electric field that exists in the shock rest frame, are responsible for the acceleration of energetic (kinetic energy  $\gg$  mean thermal energy) charged particles in shock waves.

### Shock Drift Mechanism

In the rest frame of nonparallel ( $\psi \neq 0^\circ$ , see Figure 2-7) magnetosonic shock waves, there exists a  $\vec{V} \times \vec{B}$  electric field  $\vec{E}_t$  due to the motion of the upstream and downstream magnetized plasma. For the shock geometry in Figure 2-7(b),  $\vec{E}_t = -\hat{y} |\vec{V}_{ss}| |\vec{B}_1| \sin \psi_1$ , where  $\vec{V}_{ss}$  is the shock velocity in the upstream plasma rest frame. As pointed out by Chen and Armstrong (1975; see also Hudson, 1965), the  $\nabla \vec{B}$  guiding center drift direction of ions (electrons) interacting with fast mode shock waves is parallel (antiparallel) to  $\vec{E}_t$ . Hence,  $\nabla \vec{B}$  at the shock front drives a current  $\vec{J}$ , and the particles comprising that current experience a  $\vec{J} \cdot \vec{E}_t$  energy change.

The magnitude of  $\vec{B}$  downstream of slow-mode shock waves is smaller than the magnitude of  $\vec{B}$  upstream, the reverse of the fast-mode shock case. Because of this difference, Pesses et al. (1982) pointed out that the current driven by the  $\nabla |\vec{B}|$  in slow-mode shock waves is antiparallel to  $\vec{E}_t$ . Hence, in the rest frame of slow-mode shocks, particles comprising the current  $\vec{J}$  experience a  $\vec{J} \cdot \vec{E}_t$  energy loss.

There is some controversy on whether shock-drift acceleration is a Fermi (Fermi, 1949) process (e.g., Pesses et al., 1982; Lee and Fisk, 1982). Fermi acceleration is due to the interaction of the particle's gyration velocity with the curl of  $\vec{E}$  produced by a time-varying magnetic field (Northrop, 1963). As viewed from the shock frame, the shock drift acceleration results when the particle's cross-field drift velocity has a component parallel to the d.c. component of an electric field. This suggests that the mechanism is not a Fermi process.



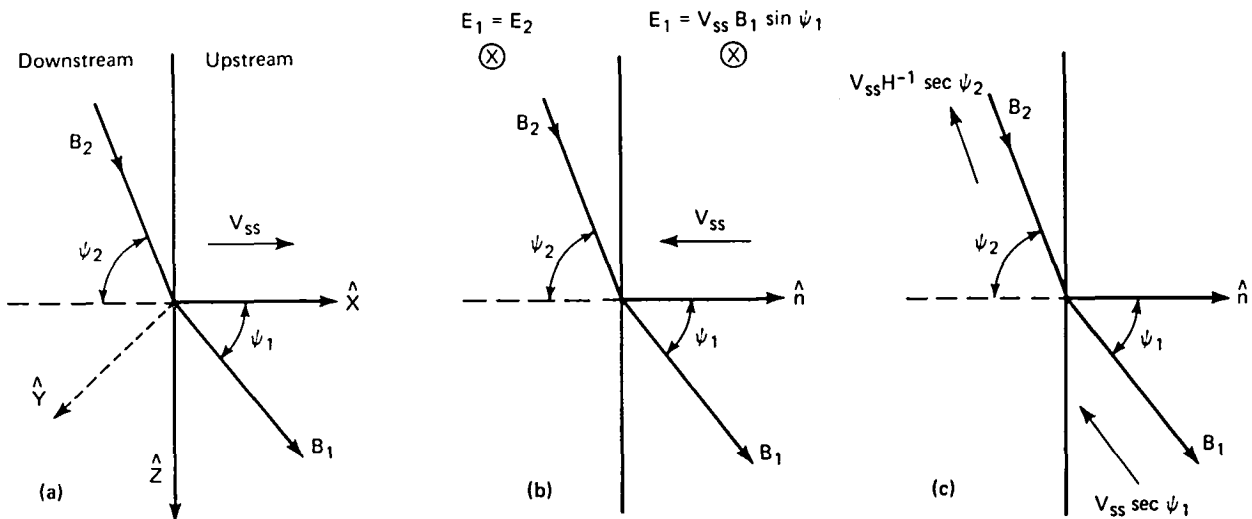


Figure 2-7. Magnetic field and flow configuration in a shock as seen from the (a) upstream and downstream plasma rest frame, (b) shock rest frame, and (c)  $E = 0$  rest frame.

The shock drift mechanism has been developed in some detail in connection with propagating quasi-perpendicular interplanetary shocks (e.g., Armstrong et al., 1977), where the spectra, pitch angle and spatial distributions, and duration have been successfully modeled and compared with the observations (Pesses et al., 1982; Decker, 1983) at higher ( $\lesssim 30$  keV) energies. This mechanism is thought to be responsible for shock-spike ion events at interplanetary traveling shocks, and may operate for energetic particles associated with corotating interaction regions (CIRs) extending to several AU in the heliosphere (Pesses et al., 1981). Application of this mechanism to astrophysical objects has been suggested (Sarris, 1975; Pesses, 1979), but has not been investigated in great detail.

### Compression Mechanism

The effect of the difference in the plasma bulk flow velocity across the shock front upon particle acceleration was pointed out independently by Axford et al. (1977), Bell (1978), and Blandford and Ostriker (1978). In the shock rest frame, particles that diffuse back and forth across the shock are accelerated by reflection off approaching upstream scattering centers and decelerated by reflection off receding downstream scattering centers. The scattering centers are convected by the bulk plasma motion, with the approaching scattering centers moving faster than the receding ones. Therefore, particles gain a net amount of energy by being effectively compressed between upstream and downstream scattering centers. This compression mechanism is a first-order (as opposed to stochastic) Fermi process, and is physically distinct from the shock drift mechanism.

As evident from the above, the shock-drift mechanism is scatter-free; and particles gain energy in a single shock encounter (one encounter involves many shock crossings). This condition obtains primarily in quasi-perpendicular shocks, and can enhance the particle energy by about an order of magnitude (Pesses et al., 1982). The compression mechanism is, by definition, a multiple encounter one, and is in principle more efficient than stochastic acceleration because it derives energy directly from the compression of the flow at the shock. For this mechanism to be effective, however, there must be adequate particle scattering both upstream and downstream of the shock; for example, adequate scattering may be provided downstream by the shock-generated turbulence, and upstream by the shock-accelerated particles themselves (cf., Bell, 1978; Lee, 1982). These conditions obtain primarily in quasi-parallel shocks, and have been modeled in the diffusive approximation for the Earth's bowshock (e.g., Lee, 1982) and for astrophysical shocks (e.g., Blandford and Ostriker, 1980).

The diffusive mechanism has been investigated most extensively in the quasi-parallel shock environment upstream of the Earth's bowshock, utilizing the wealth of new information provided by the ISEE-1, -2, -3 spacecraft (see the overview of observations included in J. Geophys. Res. 86, June 1981). A self-consistent model based on earlier work of Teresawa (1979) and Eichler (1981) has been proposed by Lee (1982), which succeeds in explaining most of the observations of the "diffuse" ion distributions and associated hydromagnetic waves, including wave and ion spectra, ion anisotropies, and the decrease in ion and wave intensity with distance upstream of the shock. Diffuse shock acceleration is also able to account for the nearly isotropic subclass of ESP events generally observed in association with quasi-parallel interplanetary traveling shocks (Forman, 1981; Scholer et al., 1983); it must, however, be noted that most observations of shock-associated ion enhancements are for quasi-perpendicular shocks (Pesses et al., 1982) even though many interplanetary quasi-parallel shocks have been observed (Sarris and Van Allen, 1974). Lee (1983) has applied the self-consistent model to quasi-parallel events with some success, and has found hydromagnetic wave power spectra in qualitative agreement with observations (cf., Tsurutani et al., 1983). It should be noted, however, that the relative importance of shock drift and compression to the observed enhancements at oblique shocks depends on the relative contributions of solar flare and solar wind ions to the seed particle populations, and can only be clarified by extensive surveys of many events and detailed observations and studies of accelerated ion composition. Thus, the issue of definitive observational tests on *in situ* acceleration vs. escape of already accelerated ions from the magnetosphere/magnetosheath upstream into the foreshock region (Edminston et al., 1981) has yet to be settled. Evidence that energetic oxygen and sulfur ions do escape from inside the magnetosphere of Jupiter (Zwickl et al., 1980, 1981) suggests that this process is clearly operating, and must be considered for the case of the Earth as well.

Finally, substantial work on this mechanism has been done for astrophysical shocks (see in particular Arons et al., 1979), although observational tests for the applicability of this process are, by necessity, difficult to come by.

### Lower Hybrid Wave Acceleration at Shocks

Cross-field currents driven by  $\nabla \vec{B}$ ,  $\nabla n$ ,  $\nabla T$ ,  $\vec{E} \times \vec{B}$  drifts and ion reflection in a collisionless shock provide an additional source of free energy for particle acceleration in a collisionless shock. As an example, consider the  $\vec{E} \times \vec{B}$  drift, which arises because of charge separation between electrons and ions in the shock transition: electrons entering the shock transition see an increase of magnetic field, and consequently experience an electric field in the shock plane. This electric field results in the acceleration of electrons, but since the ion gyroradius is larger than the shock transition, the ions are unaffected (cf., Wu, 1982). Lampe and Papadopoulos (1977) have shown that, under such circumstances, reasonable choices for the shock parameters lead to excitation of the lower hybrid drift instability at the shock front. Lower hybrid waves may then stochastically accelerate electrons in the tail of the local particle distribution (see Section on "Stochastic Acceleration" above); in particular, electron acceleration much above the phase speed

of the linearly-unstable modes can occur, as has been observed in numerical simulation studies by Tanaka and Papadopoulos (1983) and Rowland et al. (1983).

## COHERENT WAVE ACCELERATION

Although wave-particle interactions leading to particle acceleration are usually discussed in the framework of stochastic acceleration theory (e.g., "Stochastic Acceleration"), this need not be the case; in particular, the consequences of resonant wave-particle interactions have also been extensively investigated, principally in the astrophysical context of particle acceleration in pulsar magnetospheres (Gunn and Ostiker, 1969; Ostriker and Gunn, 1969; Kulsrud et al., 1972; Ferrari and Trussoni, 1975; see Kennel et al., 1979 and Arons, 1979 for further references). In any case, the question is whether effective particle acceleration can result from the trapping of particles by large amplitude waves in a plasma.

The earliest relevant work in the pulsar domain is that of Gunn and Ostriker (1969), who considered the problem of following the motion of a single test particle trapped in large amplitude electromagnetic waves; such waves were hypothesized to surround pulsars (viz., Pacini, 1967). These early calculations only considered the vacuum wave solutions; but it was soon evident that the relevant problem was one in which the effects of ambient plasmas were included. Calculations of plasma effects indeed followed soon after (based on early work of Akhiezer and Polovin (1956) on large amplitude plasma waves), including the propagation of very strong electromagnetic waves in plasmas (Max and Perkins, 1971; Max, 1973; Kennel and Pellat, 1976; Asseo et al., 1978), as well as the consequent resonant coupling between these waves and ambient particles to yield relativistic particles (cf., Kennel et al., 1973; Asseo et al., 1975); as an interesting aside, these calculations are of course quite relevant to laboratory laser-target interactions (cf., Max, 1973). The basic conclusion was that the efficiency of particle acceleration to relativistic energies was vastly reduced from the early optimistic estimates based on the vacuum wave case (e.g., Gunn and Ostriker, 1969) by the presence of ambient plasma (see Arons, 1979, and Kennel et al., 1979, for reviews and further references); and that the interaction between very large amplitude waves and particles in the magnetospheres of radio pulsars (i.e., pulsars above the pair production threshold) is far more likely to result in relativistic hot plasma (i.e., heating) than in the production of high energy cosmic rays. It should also be noted that radiation effects further limit the ultimate particle energy gain even in the vacuum limit (cf., Kulsrud, 1972; Ferrari and Trussoni, 1974). Nevertheless, under circumstances for which the magnetospheric plasma remains very tenuous such as in the magnetospheres of the presumably numerous "defunct" radio pulsars, for which pair production may be neglected), large amplitude wave acceleration may still be of considerable relevance (cf., Leboeuf et al., 1982).

In the solar-terrestrial context, it is extremely unlikely that one would encounter the relativistic wave amplitudes considered in the pulsar problem; but it is nevertheless interesting to ask whether resonant wave-particle interactions can lead to particle acceleration in the far more modest physical circumstances prevalent in the heliosphere (see, for example, Gary et al., 1968). One possibility recently reconsidered is the resonant coupling between cyclotron waves and ambient charges (for an early reference, see, for example, Kolomenskii and Lebedev, 1963); in this interaction, the relativistic cyclotron frequency and the wave phase change in such a way that the resonance between particles and the wave is maintained in an uniform magnetic field. In order for this process to be of interest, one requires sufficient wave intensity; and, indeed, it has been suggested that appropriately intense, polarized radio sources do exist within the heliosphere, including auroral kilometric radiation and intense solar microwave bursts (associated with flares; cf., Slottje, 1978). As an example of the latter case, Sprangle and Vlahos (1983) have considered acceleration of secondary electrons on "open" field lines in the vicinity of a flaring loop, within which loss-cone anisotropy of electrons leads to an intense radio burst which provides the requisite electromagnetic wave field. This research area is at yet relatively unexplored, and it is

likely that further applications will arise (particularly in the context of the magnetospheres of the major planets).

## INJECTION

A very important question in all particle acceleration theories is that of injection. Rather than discuss the injection problem in its full generality, we focus on the stochastic acceleration case (Section on "Stochastic Acceleration"); and, in order to best exemplify the problem, we further narrow our focus in the following to the solar flare context, in which we largely follow Forman et al. (1984). Nevertheless, it must be remembered that similar considerations apply to the other acceleration processes discussed above.

We first note that the basic concept of stochastic acceleration assumed that the energy changes are small compared with the particle energy, and that therefore the particle velocity must be greater than  $\delta V$ . Furthermore, for resonant scattering, ions must have  $V > V_A$  in order to scatter from (right-hand polarized) Alfvén waves, and electrons must have  $V > 43 V_A$  in order to scatter from whistlers (Melrose 1974). An additional injection condition is set by the requirement that the systematic acceleration rate due to diffusion in momentum space be larger than the ionization and Coulomb energy loss rates of the particles. The systematic acceleration rate is (e.g., Ramaty, 1979)

$$\left(\frac{dp}{dt}\right)_{\text{acc}} = (Vp^2)^{-1} \frac{\partial}{\partial p} (Vp^2 D_{pp}), \quad (14)$$

and for  $D_{pp}$  from Equation (9) above

$$\left(\frac{dE}{dt}\right)_{\text{acc}} = \frac{4}{3} \alpha \left(E(E+2Mc^2)\right)^{1/2}, \quad (15)$$

Here  $M$  and  $E$  are the proton mass and kinetic energy per nucleon for nuclei, and the electron mass and kinetic energy for electrons, respectively.

Energy loss rates due to ionization in a neutral medium and Coulomb losses in a fully ionized medium were summarized by Ramaty (1979). These loss rates, together with the systematic acceleration rates for nuclei and electrons (15), are given by Forman et al., (1984; Figures 6 and 7) for neutral and ionized media, respectively. Now, particles can be accelerated only if the rate of systematic energy gain exceeds the rate of energy loss: depending on the ratio of the acceleration efficiency,  $\alpha$ , to the ambient density,  $N$ , the energy gain curve may or may not intersect the energy loss curve. In the former case, an injection mechanism is required which preaccelerates the particles to  $E \sim E_0$ , where  $E_0$  is the energy where the two curves intersect.

The values of  $\alpha/n$  were chosen by Forman et al. (1984) such that at  $E_0$  for electrons is around 0.1 MeV, an energy at which there seems to be a transition from first phase to second phase acceleration (see, however, Section on "Solar Flares"). For protons, however, these  $\alpha/n$ 's are such that the systematic gain is larger than the loss at all energies, and so in this case it is possible to accelerate ambient particles directly (without need for a preceding injection stage). However, stochastic acceleration still requires that the particle velocities be larger than  $\delta V$ . Assuming that

$\delta V$  is of the order of the Alfvén velocity, the particles must therefore be accelerated either from a high  $\beta$  plasma ( $\beta$  = particle pressure/magnetic pressure), or possibly from the nonthermal tail of a low- $\beta$  plasma; indeed, such nonthermal tails are observed in the solar wind (Ogilvie and Scudder, 1978; Scudder and Olbert, 1979).

Finally, we note that the solar energetic particle abundances and their variability must be strongly dependent on the injection process; these injection questions have not yet been fully investigated, and they remain outstanding problems for future research (see Section IV below).

#### IV. THE REMAINING QUESTIONS

As is evident from the foregoing, the subject of particle acceleration has made enormous strides since its beginnings in cosmic ray research; most gratifying is perhaps the definitive identification of at least one particle acceleration mechanism which is both well-observed and diagnosed, and theoretically well-understood (e.g., particle acceleration at quasi-perpendicular shocks). It is however also clear that much is still to be understood; and in this final section, we wish to summarize (in essentially outline format) those major points which seem to still be outstanding research problems. This section follows the organization of Sections II and III above: we begin by discussing phenomenological issues, and then focus on outstanding theoretical (process-oriented) problems.

##### SOLAR PARTICLE ACCELERATION

The classic difficulty of solar and extra-solar cosmic ray physics is of course that (as alluded to previously) the level of detail obtained by remote sensing cannot compare -- in principle -- with that obtained by in situ measurements. For this reason, theoretical modeling can easily outstrip the ability of observations to meaningfully constrain the model. At this point, however, it seems that theoretical models can now be called upon to supply details which are (or will soon be) accessible to measurement: missing elements include not only the specific identification of the "accelerator" (viz., Fermi acceleration) and specification of its location(s), but also calculation of spectra, time profiles, and charge/mass ratio dependences, and prediction of detailed diagnostic features (such as, for example, the microwave spectrum past the self-absorption peak) within the context of such specific models. From the observational point-of-view, the missing elements include both charge state and particle spectra for escaping particle populations (for which there are presently very few measurements, especially for events seen by remote sensing instruments, such as those onboard SMM) in both "normal" and anomalous composition SEP events (e.g., He<sup>3</sup>-rich); and definitive resolution of the two-step" acceleration process problem -- is it necessary to invoke acceleration models which involve two acceleration stages? The question of localization of the acceleration and interaction sites will depend upon observations of X-rays and  $\gamma$ -rays with high angular resolution instruments (such as a pinhole occulter at high X-ray energies), or by means of multiple spacecraft observations (cf., Hurley et al., 1983).

##### INTERPLANETARY PARTICLE ACCELERATION

Here the major need at present is on the observational side. Events with associated very energetic (10-100 keV) electrons are relatively rare; hence, detailed IPM measurements of this component are very few in number. Furthermore, there is a general need for measurements of 1 AU particle events and their identification with solar transients (see paragraph on "Solar Particle Acceleration" above). What is the seed population of energetic ions (>30 keV) associated with

ESP events? Here, what seems to be required is time-resolved observations of several distinct ion particle distributions (as will be done by the ISPM spacecraft) and their anisotropies. Measurements of low ionization charge states and ion composition, as well as high mass resolution observations, are also of great interest; such measurements are particularly relevant to understanding the process(es) leading to the Anomalous Component (in particular, of the anomalous species, such as oxygen) and acceleration in corotating streams, and to resolving details of the acceleration process in the vicinity of corotating streams (viz., the importance of curved geometry in these shocks).

## **MAGNETOSPHERIC PARTICLE ACCELERATION**

Although the terrestrial magnetosphere is remarkably well diagnosed (at least by astrophysical standards), a rather significant number of problems remain to be resolved.

### **Auroral Particle Acceleration**

The most outstanding problem in auroral particle acceleration is the mechanism for the generation of the electric potentials across magnetic L shells (whose resolution may require kinetic models, although it may prove to be impractical to thereby model the entire magnetosphere). Other important problems include the question of whether the cumulative potential drop across multiple micro-double layers provides a significant fraction of the total auroral potential drop (as well as the nature of the process which allows localization of the potential drop -- micro-double layers and/or anomalous resistivity); the mechanism underlying the generation of ion conics above the auroral ionosphere; and the nature of the relationship between the cross-tail convection potential, the upward flux of ions and the field-aligned potential drop over the polar cap.

### **Magnetotail Particle Acceleration**

The outstanding problems in this domain include the process for thermalization of ions in the magnetotail as they eventually merge with the plasma sheet population; the heating of electrons that form part of the plasma sheet (i.e., is the modified two-stream instability, excited by cross-field streaming of ions in the presence of large magnetic field gradients, involved?); and, finally, the origin of the highly energetic electron and ion bursts in the Earth's magnetotail (which is tied to the nature of substorms, a major research problem in its own right, reviewed in a companion chapter in this volume).

### **Bowshock Particle Acceleration**

Of all the magnetospheric particle acceleration problems, those related to the bowshock have gained the most to date from the comparative study of distinct planetary magnetospheres within the heliosphere. For example, the source of the seed population has long been an outstanding problem; and the recent Jovian observations of sulphur upstream of the bowshock (Section II) indicate that at least some of the seed material comes from the magnetosphere itself (rather than from the solar wind). It however remains to be shown in detail what fraction of the accelerated population is of magnetospheric origin as a function of energy. A major remaining theoretical issue concerns the further elucidation of quasi-parallel shock acceleration (see below).

## EXTRASOLAR PARTICLE ACCELERATION

The major unsolved problems here concern the nature of the energy spectrum and composition above  $\approx 10^{15}$  GeV, as well as at very low energies (where the modulation by the heliosphere -- both by contamination via "local" acceleration and by propagation effects in the solar wind -- serve to mask the true nature of the extra-solar cosmic rays). Although (stellar wind and/or SNR) shocks seen to presently account for many of the observed cosmic ray properties, there remains the question of whether other acceleration sites (and related processes) may still be required; this question is particularly germane if substantial injection of a superthermal particle population is required, and is also relevant to the problem of the origin of extremely energetic cosmic rays (whose origin is not easily accounted for by invoking shock acceleration models).

## THEORETICAL MODELS FOR PARTICLE ACCELERATION

On the theoretical side, a substantial number of problems remain to be resolved, both on the level of detailed numerical (particle orbit, kinetic, or fluid) simulation, and overall modeling of the acceleration site. Because much of the details have already been covered in Section III, we limit our discussion of the major outstanding problems to an outline presentation only; the details can be appreciated by returning to our previous discussion. The summary below follows, as the phenomenological discussion immediately above, the ordering scheme of the main text (Section III).

### Adiabatic Compression and Diffusion

What is the relevant diffusion coefficient and, more particularly, what is the operative diffusion mechanism? Can one design observational tests (i.e., define the requisite observational constraints) necessary for distinguishing between distinct diffusion mechanisms?

### Direct Electric Field Acceleration

How do the final energies of accelerated particles depend upon the injection conditions, and upon the magnetic geometry (including the effect of geometry on the trapping time scale); under what conditions do microinstabilities set in, and to what extent do these provide an anomalous resistivity and how is the acceleration process affected; can particles be accelerated by the shocks which develop in the Petschek and Sonnerup reconnection models, and are there distinguishing diagnostics by which one can decide whether direct electric field, or shock, acceleration dominates in the reconnection region; can one consider the particles ejected from the reconnection region (at the local Alfvén speed) as the seed population for further acceleration (by, for example, stochastic acceleration processes)?

### Shock Acceleration

The basic problems revolve about the issues of (a) the relationship between the microstructure of shocks and the ambient plasma conditions (including, in particular, conditions upstream from the discontinuity), and (b) how the accelerated particles themselves feed back to affect the (micro-)structure of the accelerating shock. These problems are particularly acute for quasi-parallel shocks, which are not well-understood (viz., why is it, for example, that quasi-linear theory seems to "work", even though  $B/B \approx D(1) ?$ ), and for high Mach number (collisionless) shocks (including relativistic shocks, which are likely to be important to particle acceleration in the vicinity of pulsars and other collapsed objects). This area is -- just as the problem of the substorm, or reconnection, and the related particle acceleration -- a major research problem in its own right, and is the subject of a companion chapter in this volume.

As is evident from the foregoing, the problem of particle acceleration is virtually universal, and cannot be easily and neatly segregated by the classical research categories of present-day solar-terrestrial and astrophysical research. Each of the various disciplines has a clear contribution to make: the magnetospheric and interplanetary disciplines provide the in situ diagnostic opportunities which allow detailed confrontation between theories and observations; the more astrophysically-oriented disciplines provide an opportunity to test theories at their extremes, and indeed under circumstances which are at first not recognized to be relevant to the problem of particle acceleration in the deceptively-familiar near-Earth environment. It must be remembered that this distinction is based on substance, and is not at all arbitrary; whereas the in situ sensing familiar to the space plasma physicist is, in principle, denied to the solar physicist and astrophysicist, the wide range of conditions under which particle acceleration is known to occur (as revealed by remote sensing, principally via observations of the photons which result from the interaction between accelerated particles and ambient matter and fields) is not accessible to the space plasma physicist. We are, however, fortunate to be experiencing a time of dynamic interaction between these two points of view; as further theoretical modeling and new diagnostics are designed and applied, we are very likely to see much more successful interactions of the kind we have already witnessed.

## V. REFERENCES

- Achterberg, A., On the nature of small amplitude fermi acceleration, Astron. Astrophys., **97**, 259, 1981.
- Akasofu, S. -I., in The Physics of Magnetospheric Substorms, Chap. 6, D. Reidel, Dordrecht, Holland, 1977.
- Akhiezer, A. I., and R. V. Polovin, Theory of wave motion of an electron plasma, Soviet Phys. - JETP, **3**, 696, 1956.
- Alfvén, H., and C. G. Fälthammar, Cosmic Electrodynamics, Oxford, NY, 1963.
- Armstrong, T. P., and S. M. Krimigis, Interplanetary acceleration of relativistic electrons observed with IMP-7, J. Geophys. Res., **81**, 677, 1976.
- Armstrong, T. P., B. Chen, E. T. Sarris, and S. M. Krimigis, Acceleration and modulation of electrons and ions by propagating interplanetary shocks, in Study of Traveling Interplanetary Phenomena, M. A. Shea, D. F. Smart, and S. T. Wei, editors, p. 367, D. Reidel, Hingham, MA, 1977.
- Arons, J., Some problems of pulsar physics, Space Sci. Rev., **24**, 437, 1979.
- Arons, J., C. Max, and C. McKee, Particle acceleration mechanisms in astrophysics, editors, AIP No. 56, NY, 1979.



- Asbridge, J. R., S. J. Bame, and I. B. Strong, Outward flow of protons from the Earth's bowshock, J. Geophys. Res., **73**, 5577, 1968.
- Ashour-Abdalla, M., H. Okuda, and C. Z. Cheng, Acceleration of heavy ions on auroral field lines, Geophys. Res. Lett., **8**, 795, 1981.
- Asseo, E., C. F. Kennel, and R. Pellat, Flux limit on cosmic ray acceleration by strong spherical pulsar waves, Astron. Astrophys., **44**, 31, 1975.
- Asseo, E., C. F. Kennel, and R. Pellat, Synchron-compton radiation damping of relativistically strong linearly polarized plasma waves, Astron. Astrophys., **65**, 401, 1978.
- Axford, W. I., Acceleration of cosmic rays by shock waves, in Plasma Astrophysics, T. D. Guyenne and G. Levy, editors, ESA SP-161, p. 427, 1981.
- Axford, W. I., E. Leer, and G. Skadron, The acceleration of cosmic rays by shock waves, Proc. 15th Int. Conf. Cosmic Rays, **11**, 132, 1977.
- Baker, D. N., R. D. Belian, P. R. Higbie, and E. W. Hones, Jr., High-energy magnetospheric protons and their dependence on geomagnetic and interplanetary conditions, J. Geophys. Res., **84**, 7138, 1979.
- Barnes, C. W., and J. A. Simpson, Evidence for interplanetary acceleration of nucleons in corotating interaction regions, Astrophys. J. Lett., **210**, L91, 1976.
- Belian, R. D., D. N. Baker, E. W. Hones, Jr., P. R. Higbie, S. J. Bame, and J. R. Asbridge, Timing of energetic proton enhancements relative to magnetospheric substorm activity and its implication for substorm theories, J. Geophys. Res., **86**, 1415, 1981.
- Bell, A. R., The acceleration of cosmic rays in shock fronts, 1, M. N. R. A. S., **182**, 147, 1978.
- Benford, G., Turbulent resistive heating of solar coronal arches, Astrophys. J., **269**, 690, 1983.
- Blandford, R. D., and J. P. Ostriker, Shock acceleration of cosmic rays, Astrophys. J. Lett., **221**, L29, 1978.
- Blandford, R. D., and J. P. Ostriker, Supernova shock acceleration of cosmic rays in the galaxy, Astrophys. J., **237**, 793, 1980.
- Block, L. P., Potential double layers in the ionosphere, Cosm. Electrodyn., **3**, 349, 1972.
- Bohmer, H., and S. Fornaca, Experiments on non-linear effects of strong ion-cyclotron wave turbulence, J. Geophys. Res., **84**, 5234, 1979.
- Borovsky, J. E., and G. Joyce, Numerically simulated two-dimensional auroral double layers, J. Geophys. Res., **83**, 3116, 1983.
- Brown, J. C., The deduction of energy spectra of non-thermal electrons in flares from the observed dynamic spectra of hard X-ray bursts, Solar Phys., **18**, 489, 1971.
- Brown, J. C., in Solar Y-ray, X-ray and EUV Radiation, IAU Symposium No. 68, Buenos Aires, June 11-14, 1974, D. Reidel, Dordrecht, Holland, 1975.

- Blanov, S. V., The energy spectrum of particles accelerated near a singular magnetic field line, Sov. Astron. Lett., **6**, 206, 1980.
- Bulanov, S. V., and P. V. Sasorov, Energy spectrum of particles accelerated in the neighborhood of a line of zero magnetic field, Sov. Astron., **19**, 464, 1976.
- Burke, J. R., and D. Layzer, On the acceleration of ultrarelativistic electrons and nuclei in nonthermal radio sources, Astrophys. J., **157**, 1169, 1969.
- Carlqvist, P., Some theoretical aspects of electrostatic double layers in Wave Instabilities in Space Plasmas, P. J. Palmadesso and K. Papadopoulos, editors, D. Reidel, Dordrecht, p. 83, 1979.
- Casse, M., and J. A. Paul, Local gamma rays and cosmic ray acceleration by supersonic stellar winds, Astrophys. J., **237**, 236, 1980.
- Cesarsky, C. J., Cosmic ray confinement in the galaxy, Ann. Rev. Astron. Astrophys., **18**, 289, 1980.
- Chang, T., and B. Coppi, Lower hybrid acceleration and ion evolution in the superauroral region, Geophys. Res. Lett., **8**, 1253, 1981.
- Chen, G., and T. P. Armstrong, Acceleration of charged particles in oblique MHD shocks, Proc. 14th Intl. Course Ray Conf., **15**, 1814, 1975.
- Chiu, Y. T., and J. M. Cornwall, Electrostatic model of a quiet auroral arc, J. Geophys. Res., **45**, 543, 1980.
- Chiu, Y. T., J. M. Cornwall, J. F. Fennell, D. J. Gorney, and P. F. Mizera, Auroral plasmas in the evening sector: satellite observations and theoretical interpretations, Space Sci. Rev., **35**, 211, 1983.
- Chupp, E. L., Solar energetic photon transients, in Y-ray Transients and Related Astrophysical Phenomena, New York, AIP, p. 363, 1982.
- Chupp, E. L., High energy particle acceleration in solar flares: observational aspects, in Recent Advances in the Understanding of Solar Flares, Y. Uchida, S. R. Kane, H. S. Hudson, and K. Tanaka, editors, Solar Phys., in press, 1983.
- Chupp, E. L., and D. J. Forrest, Informal report submitted to SMM and STPW Workshops, 1983.
- Chupp, E. L., et al., Observation of the 2.223 MeV Gamma-ray line on the SMM satellite -- The event of 1980 June 7, Astrophys. J. Lett., **244**, L171, 1981.
- Chupp, E. L., et al., A direct observation of solar neutrons following 0118 UT flare on 1980 June 21, Astrophys. J. Lett., **263**, L95, 1982.
- Colgate, S. A., A phenomenological model of solar flares, Astrophys. J., **221**, 1068, 1978.
- Collin, H. L., R. D. Sharp, E. G. Shelley, and R. G. Johnson, Some general characteristics of upflowing ion beams over the auroral zone and their relationship to auroral electrons, J. Geophys. Res., **86**, 6820, 1981.

- Craven, J. D., and L. A. Frank, Observations of angular distributions of low energy electrons over the auroral zones with Ariel 4, Proc. Roy. Soc. London A., **343**, 167, 1975.
- Decker, R. B., The formation of shock spike events at quasi-perpendicular shocks, J. Geophys. Res., **88**, in press, 1983.
- Duijveman, A., P. Hoyng, and M. E. Machado, X-ray imaging of three flares during the impulsive phase, Solar Phys., **81**, 1982.
- Dusenberry, P. B., and L. R. Lyons, Generation of ion conic distributions by downward auroral current, J. Geophys. Res., **86**, 7627, 1981.
- Eastman, T. E., L. A. Frank, and C. Y. Haug, The boundary layers as the primary transport regions of the Earth's magnetotail, Dept. of Physics and Astronomy, Univ. of Iowa, Iowa City, IO, Report 83-7, submitted to J. Geophys. Res., 1983a.
- Eastman, T. E., L. A. Frank, W. K. Peterson, and W. Lennartsson, The plasma sheet boundary layer, Dept. of Physics & Astronomy, Univ. of Iowa, IO, Report 83-5, submitted to J. Geophys. Res., 1983b.
- Edminston, J. P., C. F. Kennel, and D. Eichler, Escape of heated ions upstream of a quasi-parallel shock, Geophys. Res. Lett., **9**, 531, 1981.
- Eichler, D., Energetic particle spectra in finite shocks: the Earth's bowshock, Astrophys. J., **244**, 711, 1981.
- Evenson, P., P. Meyer, and S. Yanagita, Solar flare shocks in interplanetary space and solar flare particle events, Int. Cosmic Ray Conf. (Paris), **3**, 32, 1981.
- Fan, C. Y., G. Gloeckler, and D. Hovestadt, The composition of heavy ions in solar energetic particle events, Space Sci. Rev., in press, 1983.
- Fennell, J. F., D. J. Gorney, and P. F. Mizera, Auroral particle distributions and their relationship to auroral arcs, in AGU Monograph No. 25, S. -I. Akasofu, editor, Washington, DC, 1979.
- Fermi, E., On the origin of the cosmic radiation, Phys. Rev., **75**, 1169, 1949.
- Fermi, E., Galactic magnetic fields and the origin of cosmic radiation, Astrophys. J., **119**, 1, 1954.
- Ferrari, A., and E. Trussoni, Acceleration by oblique magnetic rotators: near fields and radiation braking, Astron. Astrophys., **36**, 267, 1974.
- Ferrari, A., and E. Trussoni, Pulsed high energy radiation from oblique magnetic rotators, Astrophys. Space Sci., **33**, 111, 1975.
- Fillius, W., D. S. Intrilligator, M. E. Pesses, and B. Tsurutani, Five MeV electrons in coincidence with an interplanetary shock, EOS, **64**, 308, 1983.
- Fisk, L. A., B. Kozlovsky, and R. Ramaty, An interpretation of the observed oxygen and nitrogen enhancements in low energy cosmic rays, Astronphys. J. Lett., **190**, L35, 1974.

- Fisk, L. A., and M. A. Lee, Shock acceleration of energetic particles in corotating interaction regions in the solar wind, Astrophys. J., **237**, 620, 1980.
- Forrest, D. J., and E. L. Chupp, Simultaneous acceleration of electrons and ions in 2 solar flares, Nature, in press, 1983.
- Forrest, D. J., et al., Evidence for impulsive ion acceleration during the 0312 UT flare of 1980 June 7, Proc. 17th Int. Cosmic Ray Conf., **10**, 5, 1981.
- Forman, M. A., Excelleration theory for 5-40 keV ions at interplanetary shocks, Advance Space Res., Vol. I, (3), **97**, 1981.
- Forman, M. A., R. Ramaty, and E. G. Zweibel, The acceleration and propagation of solar flare energetic particles, in Physics of the Sun, T. E. Holzer, D. Mihalas, P. A. Sturrock and R. K. Ulrich, editors, in press, 1984.
- Frank, L. A., and K. L. Ackerson, Observations of charged particle precipitation into the auroral zone, J. Geophys. Res., **76**, 3612-3643, 1971.
- Friedman, M., Possible mechanism for the acceleration of ions in some astrophysical phenomena, Phys. Rev., **182**, 1408, 1969.
- Galeev, A. A., Magnetospheric tail dynamics, in Magnetospheric Plasma Physics, A. Nishida, editor, D. Reidel (Japan), p. 143, 1982.
- Garcia-Munoz, N., J. A. Simpson, and J. P. Wesel, The isotopes of neon in the galactic cosmic rays, Astrophys. J. Lett., **232**, L95, 1979.
- Gary, S. P., D. Montgomery, and D. W. Swift, Particle acceleration by electrostatic waves with spatially varying phase velocity, J. Geophys. Res., **73**, 7524, 1968.
- Ghielmetti, A. G., R. G. Johnson, R. D. Sharp, and E. G. Shelley, The latitudinal, diurnal, and altitudinal distributions of upward flowing energetic ions of ionospheric origin, Geophys. Res. Lett., **5**, 59, 1978.
- Gloeckler, G., Compositions of energetic particle populations in interplanetary space, Rev. Geophys. Space Sci., **17**, 569, 1979.
- Gloeckler, G., F. M. Ipavich, C. Y. Fan, and D. Hovestadt, Post-shock spikes: a new feature of proton and alpha enhancements associated with an interplanetary shock wave, Geophys. Res. Lett., **1**, 65, 1974.
- Goertz, C. K., Jupiter's ionosphere and magnetosphere, Planet. Space Sci., **21**, 1389, 1978.
- Gorney, D. J., A. Clarke, D. Croley, J. Fennell, J. Luhmann, and P. Mizera, The distribution of ion beams and conics below 8000 km, J. Geophys. Res., **86**, 83, 1981.
- Gosling, J. T., J. R. Asbridge, S. J. Bame, G. Paschmann, and N. Schopke, Observations of two distinct populations of bowshock ions, Geophys. Res. Lett., **5**, 957, 1978.

- Gosling, J. T., J. R. Asbridge, S. J. Bame, W. C. Feldman, R. D. Zwickl, G. Paschmann, N. Schopke, and R. J. Hynds, Interplanetary ions during an energetic storm particle event: the distribution function from solar wind thermal energies to 1.6 MeV, J. Geophys. Res., **86**, 547, 1981.
- Greenspan, M. E., and E. C. Whipple, The effect of oblique double layers on particle magnetic moment and gyrophase, Yosemite Conference on Origins of Plasmas and Electric Fields in the Magnetosphere, Yosemite, CA, 1982.
- Gunn, J. E., and J. P. Ostriker, Magnetic dipole radiation from pulsars, Nature, **221**, 454, 1969.
- Gurnett, D. A., L. A. Frank, and R. P. Lepping, Plasma waves in the distant magnetotail, J. Geophys. Res., **81**, 6059, 1976.
- Guyenne, T. D., and G. Levy, Plasma Astrophysics, editors, Varenna Plasma Astrophysics Workshop, ESA SP-161, 1981.
- Hamilton, D. C., G. Gloeckler, S. M. Krimigis, and L. J. Lanzerotti, Composition of nonthermal ions in the Jovian magnetosphere, J. Geophys. Res., **86**, 8301, 1981.
- Haerendel, G., and G. Paschmann, Interaction of the solar wind with the dayside magnetosphere, in Magnetospheric Plasma Physics, A. Nishida, editor, D. Reidel (Japan), p. 49, 1982.
- Hasselmann, K., and G. Wibberenz, Scattering of charged particles by random electromagnetic fields, J. Geophys., **34**, 353, 1968.
- Harel, M., R. A. Wolf, P. H. Reiff, R. W. Spiro, W. J. Burke, F. J. Rich, and M. Smiddy, Quantitative simulation of magnetospheric substorm, I. model logic and overview, J. Geophys. Res., **86**, 2217, 1981.
- Hayashi, T., and T. Sato, Magnetic reconnection: acceleration, heating, and shock formation, J. Geophys. Res., **83**, 217, 1978.
- Heppner, J. P., M. L. Miller, M. B. Pongratz, G. M. Smith, L. C. Smith, S. B. Mende, and N. R. Nath, The cameo barium releases:  $E_{\parallel}$  fields over the polar cap, J. Geophys. Res., **86**, 3519, 1981.
- Heyvaerts, J., Particle acceleration in solar flares, in Solar Flare Magnetohydrodynamics, E. R. Priest, editor, Gordon and Breach, NY, p. 429, 1981.
- Hones, E. W., Jr., Plasma flow in the magnetotail and its implication formulation theories, in Dynamics of the Magnetosphere, S. -I. Akasofu, editor, D. Reidel, Dordrecht, Holland, p. 545, 1980.
- Hoyng, P., K. A. Marsh, H. Zirin, and B. R. Dennis, Microwave and hard X-ray imaging of a solar flare on 1980 November 5, Astrophys. J., **285**, 865, 1983.
- Hudson, P. D., Reflection of charged particles by plasma shocks, M. N. R. A. S., **131**, 23, 1965.
- Hurley, K., M. Niel, R. Talon, I. V. Estulin, and V. Ch. Dolidze, Multispacecraft X-ray observations of fine time structure in 2 solar flares, Astrophys. J., **265**, 1076, 1983.

- Johnson, R. G., Energetic ion composition in the Earth's magnetosphere, Rev. Geophys. and Space Phys., **17**, 696, 1979.
- Jokipii, J. R., Propagation of cosmic rays in the solar wind, Rev. Geophys. Space Phys., **9**, 27, 1971.
- Jokipii, J. E., Introduction in Particle Acceleration Mechanisms in Astrophysics, J. Arons, C. Max, and C. McKee, editors, AIP No. 56, NY, 1979.
- Kahler, S., D. Spicer, Y. Uchida, and H. Zirin, Primary energy release in Solar Flares, P. A. Sturrock, editor, p. 83, Colorado Associated Univ. Press, Boulder, CO, 1980.
- Kane, S. R., et al., Impulsive phase of solar flares in Solar Flares, P. A. Sturrock, editor, Colorado Associated Univ. Press, Boulder, CO, p. 187, 1980.
- Kane, S. R., A. O. Benz, and R. A. Treumann, Electron acceleration in impulsive solar flares, Astrophys. J., **263**, 423, 1982.
- Kane, S. R., Spatial structure of high energy photon sources in solar flares, in Recent Advances in the Understanding of Solar Flares, Y. Uchida, S. R. Kane, H. S. Hudson, and K. Tanaka, editors, Solar Phys., in press, 1983.
- Kaufman, P., et al., Microwave and hard X-ray observations of a solar flare with a time resolution better than 100 msec., Solar Phys., in press, 1983.
- Kennel, C. F., and M. Ashour-Abdalla, Electrostatic waves and strong diffusion of magnetospheric electrons, Ch. 5 in Magnetospheric Plasma Physics, A. Nishida, editor, D. Reidel, Dordrecht, 1982.
- Kennel, C. F., and R. Pellat, Relativistic non-linear plasma waves in a magnetic field, J. Plasma Phys., **15**, 335, 1976.
- Kennel, C. F., G. Schmidt, and T. Wilcox, Cosmic-ray generation by pulsars, Phys. Rev. Lett., **31**, 1364, 1973.
- Kennel, C. F., F. S. Fujimura, and R. Pellat, Pulsar magnetospheres, Space Sci. Rev., **24**, 407, 1979.
- Kindel, J. M., and C. F. Kennel, Topside current instabilities, J. Geophys. Res., **26**, 3053, 1971.
- Kitner, P. M., and D. J. Gorney, A search for the plasma processes associated with the source of ion conics, J. Geophys. Res., in press, 1983.
- Kitner, P. M., M. C. Kelley, and F. S. Mozer, Electrostatic hydrogen waves near one earth radius altitude in the polar magnetosphere, Geophys. Res. Lett., **5**, 139, 1978.
- Kiplinger, A. L., B. R. Dennis, A. G. Emslie, K. J. Frost, and L. E. Orwig, Fast transients in hard X-ray solar flares, in Recent Advances in the Understanding of Solar Flares, Y. Uchida, S. R. Kane, H. S. Hudson, and K. Tanaka, editors, Solar Phys., in press, 1983.

- Klecker, B., M. Scholer, D. Hovestadt, G. Gloeckler, and F. M. Ipavich, Spectral and compositional variations of low energy ions during an energetic storm particle event, Astrophys. J., **251**, 393, 1981.
- Klumpar, D. M., Transversely accelerated ions: an ionospheric source of hot magnetospheric ions, J. Geophys. Res., **84**, 4229, 1979.
- Knight, L., Parallel electric fields, Planet. Space Sci., **2**, 741, 1975.
- Kilomenskii, A. A., and A. N. Lebedev, Self-resonant particle motion in a plane electromagnetic wave, Solar Phys., **7**, 745, 1963.
- Krimigis, S. M., and E. T. Sarris, Energetic particle bursts in the earth's magnetotail, in Dynamics of the Magnetosphere, S. -I. Akasofu, editor, D. Reidel, Dordrecht, Holland, p. 599, 1980.
- Krimigis, S. M., J. F. Carbary, E. P. Keath, T. P. Armstrong, L. J. Lanzerotti, and G. Gloeckler, General characteristics of hot plasma and energetic particles in the saturnian magnetosphere, J. Geophys. Res., **88**, in press, October, 1983.
- Kulsrud, R. M., The effect of near fields on the pulsar acceleration of particles, Astrophys. J. Lett., **174**, L25, 1972.
- Kulsrud, R., Stochastic acceleration by hydromagnetic turbulence, in Particle Acceleration Mechanisms in Astrophysics, J. Arons, C. Max, and C. McKee, editors, AIP, p. 13, 1979.
- Kulsrud, R., and A. Ferrari, Quasi-linear theory of particle acceleration by hydromagnetic waves, Astrophys. Space Sci., **12**, 302, 1971.
- Kulsrud, R. M., J. P. Ostriker, and J. E. Gunn, Acceleration of cosmic rays in supernova remnants, Phys. Rev. Lett., **28**, 636, 1972.
- Kundu, M. R., and T. E. Gergely, Radio Physics of the Sun, editors, D. Reidel, Dordrecht, Holland, 1980.
- Kundu, M. R., E. J. Schmahl, T. Velusamy, and L. Vlahos, Radio imaging of solar flares using the very large array: new insights into flare processes, Astron. Astrophys., **108**, 188, 1982.
- Lampe, M., and K. Papadopoulos, Formation of fast electron tails in type II solar bursts, Astrophys. J., **212**, 886, 1977.
- Lang, A., and H. Bohmer, Electron current disruption and parallel electric fields associated with electrostatic ion-cyclotron waves, J. Geophys. Res., **88**, 5564, 1983.
- Lanzerotti, L. J., and A. Wolfe, Particle diffusion in the geomagnetosphere: comparison estimates from magnetic and electric field fluctuations, J. Geophys. Res., **85**, 1346, 1980.
- Leboeuf, J. N., M. Ashour-Abdalla, T. Tajima, C. F. Kennel, F. V. Coroniti, and J. M. Dawson, Ultrarelativistic waves in overdense electron-position plasmas, Phys. Rev. A, **25**, 1023, 1982.
- Lee, M. A., Coupled hydromagnetic wave excitation and ion acceleration upstream of the earth's bowshock, J. Geophys. Res., **87**, 5063, 1982.

- Lee, M. A., Coupled hydromagnetic wave excitation and ion acceleration at interplanetary travelling shocks, J. Geophys. Res., in press, 1983.
- Lee, M. A., and L. A. Fisk, Shock acceleration of energetic particles in the heliosphere, Space Sci. Rev., **32**, 205, 1982.
- Lemaire, J., and M. Scherer, Ionosphere-plasma sheet field-aligned currents and parallel electric fields, Planet. Space Sci., **22**, 1485, 1971.
- Levine, R. H., Acceleration of thermal particles in collapsing magnetic regions, Astrophys. J., **190**, 447, 1974.
- Lin, R. P., C. I. Meng, and K. A. Anderson, 30 to 100 keV protons upstream from the earth's bowshock, J. Geophys. Res., **79**, 489, 1974.
- Luhmann, J. G., A quasi-linear kinetic equation for cosmic rays in the interplanetary medium, J. Geophys. Res., **81**, 2089, 1976.
- Lyons, L. R., Generation of large-scale regions of auroral currents, electric potentials and precipitation by the divergence of the convection electric field, J. Geophys. Res., **85**, 17, 1980.
- Lyons, L. R., and T. W. Speiser, Evidence for current sheet acceleration in the geomagnetic tail, J. Geophys. Res., **87**, 2276, 1982.
- Lysak, R. L., M. K. Hudson, and M. Temerin, Ion heating by strong electrostatic ion cyclotron turbulence, J. Geophys. Res., Vol. 85, **678**, 1980.
- Manheimer, W., and J. P. Boris, Marginal stability analysis - a simpler approach to anomalous transport in plasmas, Comments Plasma Phys. Contr. Fusion, **3**, 15, 1977.
- Marsh, K. A., and G. T. Hurford, Two dimensional VLA maps of solar bursts at 15 and 23 GHz with arcsec resolution, Astrophys. J. Lett., **240**, L111, 1980.
- Marshall, F. E., and E. C. Stone, Persistent sunward flow of 1.6 MeV protons at 1 AU, Geophys. Res. Lett., **4**, 57, 1977.
- Max, C. E., Steady-state solutions for relativistically strong electromagnetic waves in plasmas, Phys. Fluids, **16**, 1277, 1973.
- Max, C. E., and F. Perkins, Strong electromagnetic waves in overdense plasmas, Phys. Rev. Lett., **27**, 1342, 1971.
- McDonald, F. B., The acceleration of particles on the sun and in the heliosphere, in 17th Int. Cosmic Ray Conf. (Paris), **13**, 199, 1981.
- McDonald, F. B., B. T. Teegarden, J. H. Trainor, and R. R. von Rosenvinge, The interplanetary acceleration of energetic nucleons, Astrophys. J. Lett., **203**, L149, 1976.
- McGuire, R. E., T. T. Zon Rosenvinge, and F. B. McDonald, 17th int cosmic ray coms., (Paris) Vol. 3, **65**, 1981.
- Melrose, D. B., Resonant scattering of particles and second phase acceleration in the solar corona, Solar Phys., **37**, 353, 1974.



- Melrose, D. B., Plasma Astrophysics, Gordon and Breach, NY, 1980.
- Mewaldt, R. A., E. C. Stone, and R. E. Vogt, The radial diffusion coefficient of 1.3-2.3 MeV protons in recurrent proton stream, Geophys. Res. Lett., **5**, 11, 1978.
- Mewaldt, R. A., J. D. Spalding, E. C. Stone, and R. E. Vogt, High resolution measurements of galactic cosmic ray neon, magnesium, and silicon isotopes, Astrophys. J. Lett., **235**, L95, 1980.
- Mizera, P. F., J. F. Fennell, D. R. Croley, A. L. Vampola, F. S. Mozer, R. B. Torbert, M. Temerin, R. L. Lysak, M. K. Hudson, C. A. Cattell, R. G. Johnson, R. D. Sharp, A. Ghielmetti, and P. M. Kintner, The auroral inferred from S3-3 particles and fields, J. Geophys. Res., **86**, 2329, 1981a.
- Mizera, P. F., J. F. Fennell, D. R. Croley, and D. J. Gorney, Charged particle distributions and electric field measurements from S3-3, J. Geophys. Res., **86**, 2329, 1981b.
- Mozer, F. S., C. W. Carlson, M. K. Hudson, R. B. Torbert, B. Parady, J. Yatteau, and M. C. Kelley, Observations of paired electrostatic shocks in the polar magnetosphere, Phys. Rev. Lett., **38**, 292, 1977.
- Mozer, F. S., C. A. Cattell, M. K. Hudson, R. L. Lysak, M. Temerin, and R. B. Torbert, Satellite measurements and theories of low altitude auroral particle acceleration, Space Sci. Rev., **27**, 155, 1980.
- Mullan, D. J., and R. H. Levine, Preacceleration in collapsing magnetic neutral sheets and anomalous abundance of solar flare particles, Astrophys. J. Suppl., **47**, 87, 1981.
- Northrop, T. G., The adiabatic motion of charged particles, Interscience Publishers, Wiley, NY, 1963.
- Ogilvie, K. W., and J. D. Scudder, The radial gradients and collisional properties of solar wind electrons, J. Geophys. Res., **83**, 3776, 1978.
- Ostriker, J. P., and J. E. Gunn, On the nature of pulsar, I. theory, Astrophys. J., **157**, 1395, 1969.
- Pacini, F., Energy emission from a neutron star, Nature, **216**, 567, 1967.
- Palmer, I. D., and J. R. Gosling, Shock-associated energetic proton events at large heliocentric distances, J. Geophys. Res., **83**, 2037, 1978.
- Papadopoulos, K., A review of anomalous resistivity in the ionosphere, Rev. Geophys. Space Phys., **15**, 113, 1977.
- Papadopoulos, K. and T. C. Coffey, Anomalous resistivity of the auroral plasma, J. Geophys. Res., **79**, 1558, 1974.
- Papadopoulos, K., J. Coffey, and P. Palmadesso, Stochastic acceleration of large  $M$  ions by hydrogen cyclotron waves in the magnetosphere, Geophys. Res. Lett., **7**, 1014, 1980.
- Paschmann, G., N. Sckopke, J. R. Asbridge, S. J. Bame, and J. T. Gosling, Energization of solar wind ions by reflection from the earth's bowshock, J. Geophys. Res., **85**, 4689, 1983.

- Parker, E. N., Sweet's mechanism for merging magnetic fields in conducting fluids, J. Geophys. Res., **62**, 509, 1957.
- Parker, E. N., *Interplanetary Dynamical Processes*, Wiley, NY, 1963.
- Parker, E. N., *Cosmical Magnetic Fields*, Clarendon Press, Oxford, 1979.
- Parker, E. N., and D. A. Tidman, Suprathermal particles, Phys. Rev., **111**, 1206, 1958.
- Pesses, M. E., The acceleration of energetic charged particles by interplanetary and supernova shock waves, in Particle Acceleration Mechanisms in Astrophysics, J. Arons, C. Max, and C. McKee, editors, AIP No. 56, NY, 1979.
- Pesses, M. E., B. Tsurutani, J. A. Van Allen, and E. J. Smith, Acceleration of energetic protons by interplanetary shocks, J. Geophys. Res., **84**, 7297, 1979.
- Pesses, M. E., J. R. Jokipii, and D. Eichler, Cosmic ray drift, shock wave acceleration, and the anomalous component of cosmic rays, Astrophys. J. Lett., **246**, L85, 1981.
- Pesses, M. E., R. B. Decker, and T. P. Armstrong, The acceleration of charged particles in interplanetary shock waves, Space Sci. Rev., **32**, 185, 1982.
- Petschek, H. E., Magnetic field annihilation, in Physics of Solar Flares, W. N. Hess, editor, NASA SP-50, p. 425, 1964.
- Priest, Solar Magnetohydrodynamics, Dordrecht Reidel, 1983.
- Ramaty, R., Gyrosynchrotron emission and absorption in magnetoactive plasma, Astrophys. J., **158**, 753, 1969.
- Ramaty, R., Energetic particles in solar flares, in Particle Acceleration Mechanisms in Astrophysics, J. Arons, C. Max, and C. McKee, editors, AIP No. 56, NY, p. 135, 1979.
- Ramaty, R., et al., Energetic particles in solar flares, in Solar Flares, P. A. Sturrock, editor, Colorado Associated Univ. Press, Boulder, CO, 1980.
- Ramaty, R., R. J. Murphy, B. Kozlovsky, and R. E. Lingenfelter, Gamma ray lines and neutrons from solar flares, in Recent Advances in the Understanding of Solar Flares, Y. Uchida, S. R. Kane, H. S. Hudson, and K. Tanaka, editors, Solar Phys., in press, 1983a.
- Ramaty, R., R. J. Murphy, B. Kozlovsky, and R. E. Lingenfelter, Implications of high-energy neutron observations of solar flares, Astrophys. J. Lett., in press, 1983b.
- Rowland, H. L., and P. J. Palmadesso, Anomalous resistivity due to low frequency turbulence, J. Geophys. Res., in press, 1983.
- Rowland, H. L., P. J. Palmadesso, and K. Papadopoulos, Anomalous resistivity on auroral field lines, Geophys. Res. Lett., **8**, 1257, 1981.
- Rowland, H. L., M. Tanaka, and K. Papadopoulos, Plasma eigenmodes and particle acceleration, Comments Plasma Phys., in press, 1983.

- Russell, C. T., and M. M. Hoppe, Upstream waves and particles, Space Sci. Rev., **34**, 155, 1982.
- Russell, C. T., and E. W. Greenstadt, Plasma boundaries and shocks, Rev. Geophys. Space Phys., **21**, 449, 1983.
- Sarris, E. T., On the acceleration of cosmic rays, Astrophys. Space Sci., **36**, 467, 1975.
- Sarris, E. T., S. M. Krimigis, and T. P. Armstrong, Observations of high energy protons and electrons at 35R<sub>E</sub> with IMT-7, J. Geophys. Res., Vol. 81, **2341**, 1976.
- Sarris, E. T., S. M. Krimigis, C. O. Bostrom, and T. P. Armstrong, Simultaneous multi-spacecraft observations of energetic proton bursts inside and outside the magnetosphere, J. Geophys. Res., **83**, 4289-4305, 1978.
- Sarris, E. T., and J. A. Van Allen, Effects of interplanetary shock waves on energetic charged particles, J. Geophys. Res., **79**, 4157, 1974.
- Sato, T., and H. Okuda, Numerical simulations on ion acoustic double layers, J. Geophys. Res., **86**, 3357, 1981.
- Scholer, M., and F. M. Ipavich, Energetic ions upstream of the earth's bow shock during an energetic storm particle event, J. Geophys. Res., **88**, 5715, 1983.
- Scholer, M., F. M. Ipavich, G. Gloeckler, and D. Hovestadt, Acceleration of low energy protons and alpha particles at interplanetary shock waves, J. Geophys. Res., **88**, 1977, 1983.
- Schwartz, S. J., M. F. Thomsen, and J. T. Gosling, Ions upstream of the earth's bow shock: a theoretical comparison of alternative source populations, J. Geophys. Res., **88**, 2039, 1983.
- Scudder, J. D., and S. Olbert, A theory of local and global processes which affect solar wind electrons: 2. experimental support, J. Geophys. Res., **84**, 6603, 1979.
- Sharp, R. D., R. G. Johnson, and E. G. Shelley, Observation of an ionospheric acceleration mechanism producing energetic (keV) ions primarily normal to the geomagnetic field direction, J. Geophys. Res., **83**, 3324, 1977.
- Shelley, E. G., Heavy ions in the magnetosphere, Space Sci. Rev., **23**, 465, 1979.
- Shelley, E. G., R. D. Sharp, and R. G. Johnson, He<sup>++</sup> and H<sup>+</sup> flux measurements in the dayside cusp: estimates of convection electric field, J. Geophys. Res., **82**, 2363, 1976a.
- Shelley, E. G., R. D. Sharp, and R. G. Johnson, Satellite observations of an ionospheric acceleration mechanism, Geophys. Res. Lett., **3**, 654, 1976b.
- Slottje, C., Millisecond microwave spikes in a solar flare, Nature, **275**, 520, 1978.
- Skilling, J. A., Cosmic ray streaming - I. effects of Alfvén waves on particles, M. N. R. A. S., **172**, 557, 1975.
- Smith, D. F., First phase acceleration mechanisms in solar flares, in Particle Acceleration Mechanisms in Astrophysics, J. Arons, C. Max, and C. McKee, editors, AIP, p. 155, 1979.

- Sonnerup, B.U.Ö., Acceleration of particles reflected at shock front, J. Geophys. Res., **74**, 1301, 1969.
- Sonnerup, B.U.Ö., Magnetic field reconnexion in a highly conducting incompressible fluid, J. Plasma Phys., **4**, 161, 1970.
- Sonnerup, B.U.Ö., Magnetic field reconnection and particle acceleration, in High Energy Phenomena on the Sun, R. Ramaty and F. E. Stone, editors, NASA SP-342, p. 357, 1973.
- Speiser, T. W., Particle trajectories in model current sheets, J. Geophys. Res., **70**, 4219, 1965.
- Speiser, T. W., Particle trajectories in model current sheets, 2. application to auroras using a geomagnetic tail model, J. Geophys. Res., **72**, 3919, 1967.
- Sprangle, P., and L. Vlahos, Electron cyclotron wave acceleration outside a flaring loop, Astrophys. J. Lett., in press, 1983.
- Stern, D. P., and N. F. Ness, Planetary magnetospheres, Ann. Rev. Astron. Astrophys., **20**, 139, 1982.
- Sturrock, P. A., Flare models, in Solar Flares, P. A. Sturrock, editor, Colorado Associated Univ. Press, Boulder, CO, p. 411, 1980.
- Sweet, P. A., The neutral point theory of solar flares, in Electromagnetic Phenomena in Cosmical Physics, B. Lehnert, editor, Cambridge Univ. Press, London, p. 123, 1958.
- Swift, D. W., The effect of the neutral sheet on magnetospheric dynamics, J. Geophys. Res., **82**, 1288, 1977.
- Swift, D. W., An equipotential model for auroral arcs: the theory of two-dimensional laminar electrostatic shocks, J. Geophys. Res., **84**, 6427, 1979.
- Swift, D. W., Mechanisms for auroral precipitation: a review, Rev. Geophys. Space Phys., **19**, 185, 1981.
- Swift, D. W., Numerical simulation of the interaction of the plasma sheet with the lobes of the earth's magnetotail, J. Geophys. Res., **87**, 2287, 1982.
- Swift, D. W., A two-dimensional simulation of the interaction of the plasma sheet with the lobes of the earth's magnetotail, J. Geophys. Res., **88**, 125, 1983.
- Syrovatskii, S. I., Pinch sheets and reconnection in astrophysics, Ann. Rev. Astron. Astrophys., **19**, 163, 1981.
- Tanaka, M., and K. Papadopoulos, Creation of high energy electron tails due to the modified two-stream instability, Phys. Fluids, **26**, 1697, 1983.
- Tanaka, M., C. C. Goodrich, D. Winske, and K. Papadopoulos, A source of the backstreaming ion beams in the foreshock region, J. Geophys. Res., **88**, 3046, 1983.
- Temerin, M., K. Cerny, W. Lotko, and F. S. Mozer, Observations of double layers and solitary waves in the auroral plasmas, Phys. Rev. Lett., **48**, 1175, 1982.

- Terasawa, T., Origin of 30-100 keV protons observed in the upstream region of the earth's bow shock, Planet. Space Sci., **27**, 365, 1979.
- Thorne, R. M., The importance of wave-particle interactions in the magnetosphere, in Critical Problems of Magnetospheric Physics, E. R. Dyer, editor, National Academy of Sciences, Washington, DC, p. 211, 1972.
- Thorne, R. M., Microscopic plasma processes in the Jovian magnetosphere, Ch. 12 in Physics of the Jovian Magnetosphere, A. Dessler, editor, Cambridge Univ. Press, London, 1983.
- Tsurutani, B. T., E. J. Smith, and D. E. Jones, Waves observed upstream of interplanetary shocks, J. Geophys. Res., in press, 1983.
- Tsurutani, B. T., E. J. Smith, K. R. Pyle, and J. A. Simpson, Energetic protons accelerated at corotating shocks: Pioneer 10 and 11 observations from 1 to 6 AU, J. Geophys. Res., **87**, 7389, 1982.
- Wagner, J. S., J. R. Kan, and S. -I. Akasofu, Particle dynamics in the plasma sheet, J. Geophys. Res., **85**, 891, 1979.
- Wagner, J. S., J. R. Kan, S. -I. Akasofu, T. Tajima, J. N. Lebeouf, and J. M. Dawson, A simulation study of V-potential double layers and auroral arc deformations, in Physics of Auroral Arc Formation, S. -I. Akasofu and J. R. Kan, editors, American Geophysical Union, p. 305, 1981.
- Walt, M., Radial diffusion of trapped particles and some of its consequences, Rev. Geophys. Space Phys., **9**, 11, 1971.
- Webber, W. R., and A. C. Cummings, Voyager measurements of the energy spectrum, charge composition, and long term temporal variations of the anomalous components in 1977-1982, Proc. Solar Wind V, NASA, in press, 1983.
- Wentzel, D. G., The propagation and anisotropy of cosmic rays. II. electrons, Astrophys. J., **157**, 545, 1969.
- Wenzel, K. P., R. Reinhard, and T. R. Sanderson, Interplanetary acceleration of low energy protons as observed during the 25 September 1978 shock event, Adv. Space Sci., **1(3)**, 105, 1981.
- Whalen, B. A., and P. W. Daly, Do field aligned auroral particles distributions imply acceleration by quasi-static electric fields?, J. Geophys. Res., **82**, 1525, 1977.
- Wu, C. S., Physical mechanisms for turbulent dissipation in collisionless shock waves, Space Sci. Rev., **32**, 83, 1982.
- Vasyliuna, V. M., Theoretical models of magnetic field line merging, 1, Rev. Geophys. Space Phys., **13**, 308, 1975.
- Young, D. T., Ion composition measurements in magnetospheric modeling, in Quantitative Modeling of Magnetospheric Processes, American Geophysical Union, Washington, DC, p. 340, 1979.

- Zweibel, E. G., and D. Haber, The propagation of energetic ions in magnetic loops and gamma-ray emission from solar flares, Astrophys. J., **264**, 648, 1983.
- Zwickl, R. D., S. M. Krimigis, T. P. Armstrong, and L. J. Lanzerotti, ions of Jovian origin observed by Voyager 1 and 2 in interplanetary space, Geophys. Res. Lett., **7**, 453-456, 1980.
- Zwickl, R. D., S. M. Krimigis, J. F. Carbary, E. P. Keath, T. P. Armstrong, D. C. Hamilton, and G. Gloeckler, Energetic particle events ( $>30$  keV) of jovian origin observed by Voyager 1 and 2 in interplanetary space, J. Geophys. Res., **86**, 8125-8140, 1981.

# **CHAPTER 3**

## **EVOLUTION OF SOLAR MAGNETIC FLUX**

### **WORKING GROUP MEMBERS**

J. P. Boris, Chairman  
*Naval Research Laboratory*

C. R. DeVore  
*Naval Research Laboratory*

L. Golub  
*S.A.O.*

R. F. Howard  
*Mount Wilson Observatory*

B. C. Low  
*National Center for Atmospheric Research*

N. R. Sheeley, Jr.  
*Naval Research Laboratory*

G. W. Simon  
*Air Force Geophysics Laboratory*

K. C. Tsinganos  
*Harvard Smithsonian Center for Astrophysics*

### CHAPTER 3

#### EVOLUTION OF SOLAR MAGNETIC FLUX

I.	Introduction	3-3
II.	Appearance of Magnetic Flux	3-5
III.	Dynamics of Surface Magnetic Flux	3-13
IV.	Disappearance of Magnetic Flux	3-18
V.	Summary	3-23
VI.	References	3-29



## I. INTRODUCTION

Hale discovered the association of intense magnetic fields with sunspots 75 years ago. Observation of the sunspots for hundreds of years had recorded not only the fairly regular 11-year sunspot cycles but also irregular behavior, such as the Maunder Minimum, a period of about 70 years from 1645 to 1715 when essentially no sunspots were seen. Hale's discovery, however, was the first hint of complex magnetic field activity on the Sun. Today we know that sunspots and the associated Bipolar Magnetic Regions (BMRs) appear on the Sun in large-scale patterns. They evolve, interact, migrate across the solar surface and then disappear following the sunspot cycle. Since the Sun's average dipole field reverses polarity every 11 years, a complete magnetic cycle lasts 22 years. Ground-based observations still provide most of our data base on the sunspots and magnetic activity.

In the last 25 years steadily improving observations, many of which are obtained above the atmosphere, have shown that the dynamic evolution of solar magnetic fields controls most components of solar variability affecting the geophysical environment such as disruptions of global communications. Observational and theoretical advances have given us a better view of the dynamic, sometimes violent, MHD and plasma phenomena which control sunspots and the solar transition region, heat the corona, drive the solar wind, and power flares, bright points, and spicules. Most of the phenomena investigated by the other eleven working groups depend on solar magnetic field variability to drive them or to trigger subsequent mechanisms which drive them.

There are other important reasons why a better understanding of the evolution of the solar magnetic flux is desired. Because the surface features can be observed in some detail, the Sun is an important source of information on plasma and MHD phenomena which can also occur in the laboratory, in our atmosphere, and in space. The Sun's large space scales and correspondingly longer time scales make possible observations which cannot be obtained on the Earth. The Sun is also a star and therefore provides a wealth of information for extrapolation to and interpretation of other stars.

To understand the evolution of solar magnetic flux we must answer the fundamental question: How does the solar dynamo work? Recent observations certainly influence models of the dynamo but we have not been able to deduce the subsurface flows which control the generation of magnetic flux and presumably the entire solar cycle. The universality and importance of the scientific issues have attracted some of our best scientific minds and kept them busy for the better part of a century. The recent monograph on cosmical magnetic fields by Parker (1979a) serves as an excellent basic reference. Developments in magnetoconvection, dynamo theory, and their application to the solar cycle have been reviewed by Stix (1976, 1981) and Proctor and Weiss (1982). The articles by Zwaan (1981), Gilman (1981), and Spruit (1981) also review important aspects of magnetic activity, global circulation and the magnetic dynamo.

The first modern qualitative model of the operation of the solar activity cycle was provided by Babcock (1961). He considered the evolution of an axisymmetric dipole field of concentrated flux ropes subject to differential rotation. His model was able to account for the reversal of sunspot polarities from cycle to cycle, the drift of sunspot activity from mid-latitudes toward the solar equator during the cycle, and the reversal of the global dipole field near each sunspot maximum. Leighton (1969) quantified some aspects of the model by obtaining numerical solutions of the linear dynamo equations. These solutions exhibited behavior akin to that of the large-scale magnetic fields on the Sun, including fluctuations in period and amplitude in response to fluctuations in the rate of emergence of new flux.

Perhaps the most striking new finding about magnetic activity cycles is other evidence of its intermittency: intervals of aperiodic behavior interspersed irregularly with periods of dormancy (Eddy 1978). Observations of stellar calcium II H and K emission (Vaughan 1983) indicate that stars also exhibit magnetic cycles, and that some, in particular fast rotators, show aperiodic magnetic activity. It appears that intermittency and aperiodicity may be natural consequences of nonlinear dynamos. Calculations for model dynamos with interaction between the magnetic field and differential rotation have shown a transition from periodic to quasi-periodic to chaotic behavior as parameters of the dynamo change (Cattaneo et al., 1983) and have produced a convective flow which shows both differential rotation and a magnetic cycle (Yoshimura 1978a, b, 1979).

Lacking a clear view of subphotospheric phenomena, this review concentrates on what can be deduced from observations by addressing four questions which shape our discussion in the next three sections.

- (1) How does flux emerge on the Sun?
- (2) How does magnetic flux disappear?
- (3) What is seen at the limits of resolution?
- (4) What large-scale patterns are detected?

The answer to the fundamental question, How the solar dynamo works, can only evolve from interpreting clues uncovered in research aimed at illuminating various aspects of these more "visible" questions. We discuss the Appearance of Magnetic Flux in Section II, considering both the mechanisms and the phenomena associated with flux emergence. The Dynamics of Surface Magnetic Flux are treated in Section III. Both the transport and structure of magnetic-flux-carrying elements, as seen in the chromosphere and corona, are discussed. The Disappearance of Magnetic Flux from the surface of the Sun is the subject of Section IV. In each section we first discuss some of the standard models which represent generally accepted views, and then consider observations which seem to defy explanation within the context of these standard models.

This paper is the work of one of twelve working groups convened jointly at the NASA/NSF Solar-Terrestrial Physics Workshop. The participants in the Working Group on "Evolution of Solar Magnetic Flux" are:

J.P. Boris  
C.R. DeVore  
L. Golub  
R.F. Howard  
B.C. Low  
N.R. Sheeley, Jr.  
G.W. Simon  
K.C. Tsinganos

Additional commentary and contributions were received from V. Gaizauskas, R. Rosner, A. Title, J. Toomre, and E. Zweibel. The workshop was held 6-10 June 1983 at Coolfont Conference Center, Berkeley Springs, West Virginia.

## II. APPEARANCE OF MAGNETIC FLUX

Decades of detailed solar observations lead us strongly to the conclusion that magnetic flux emerges onto the surface of the Sun from the interior. We present the current "standard" picture of magnetic flux emergence in this section. However, there is still controversy over the rate of rise of buoyant flux tubes, the mechanisms and process of emergence, and the concentration of magnetic field to the strengths observed in the flux tubes. Thus the standard picture also provides a starting point for discussing the more controversial and speculative aspects of the problem appearing in succeeding sections. In recent reviews Spruit (1981) and Proctor and Weiss (1982) consider many of the important aspects of magnetic flux tubes.

We want to know what flux tube emergence from the photosphere into the corona looks like because only the magnetic field at or near the photosphere can be observed directly. Magnetic field at greater depths is not visible because the region below the photosphere is opaque. Diagnosis of the much weaker magnetic fields above the photosphere is generally not direct but is inferred from the geometry of radiating plasma which is shaped by these coronal fields. Thus magnetograms are generally analyzed treating the photosphere as a thin surface.

Magnetic flux is observed to emerge from the solar interior in concentrated "active" regions which occupy only a small fraction of the surface. These active regions are composed of numerous magnetic flux "bundles", a term generally used interchangeably with "ropes" or "tubes". Flux bundles appear in a broad spectrum of sizes. The largest active regions contain up to  $10^{23}$  Mx of flux and have complex spot groups ( $1 \text{ Mx} = 1 \text{ Maxwell} = 1 \text{ G cm}^2 = 1 \text{ Gauss cm}^2$ ). The smallest detectable ephemeral regions have only  $10^{18}$ - $10^{19}$  Mx but appear in great numbers everywhere on the Sun. Observations of the small-scale ephemeral regions, containing less than  $\sim 10^{20}$  Mx of magnetic flux, and their often associated X-ray bright points have been possible for about one solar cycle. It is clear that these small scale emergence events are major contributors of emerging flux at nearly all phases of the solar cycle. The number of bipolar regions emerging per unit time monotonically decreases with the quantity of flux in the region and the bigger regions seem to last longer (Golub 1980).

In the standard model of flux emergence presented in the next section entitled "Models for Flux Emergence" we assume that the observed magnetic fields are produced by dynamo activity in the Sun. The periodic reversal of the polar fields and the total amount of flux appearing each cycle seem to argue against the concept of a continual shedding of primordial flux which became trapped when the Sun was formed. Then, in the section on "Unexplained Observations" we discuss observations which do not easily fit into the standard picture, with possible implications for future modification of the models.

### MODELS FOR FLUX EMERGENCE

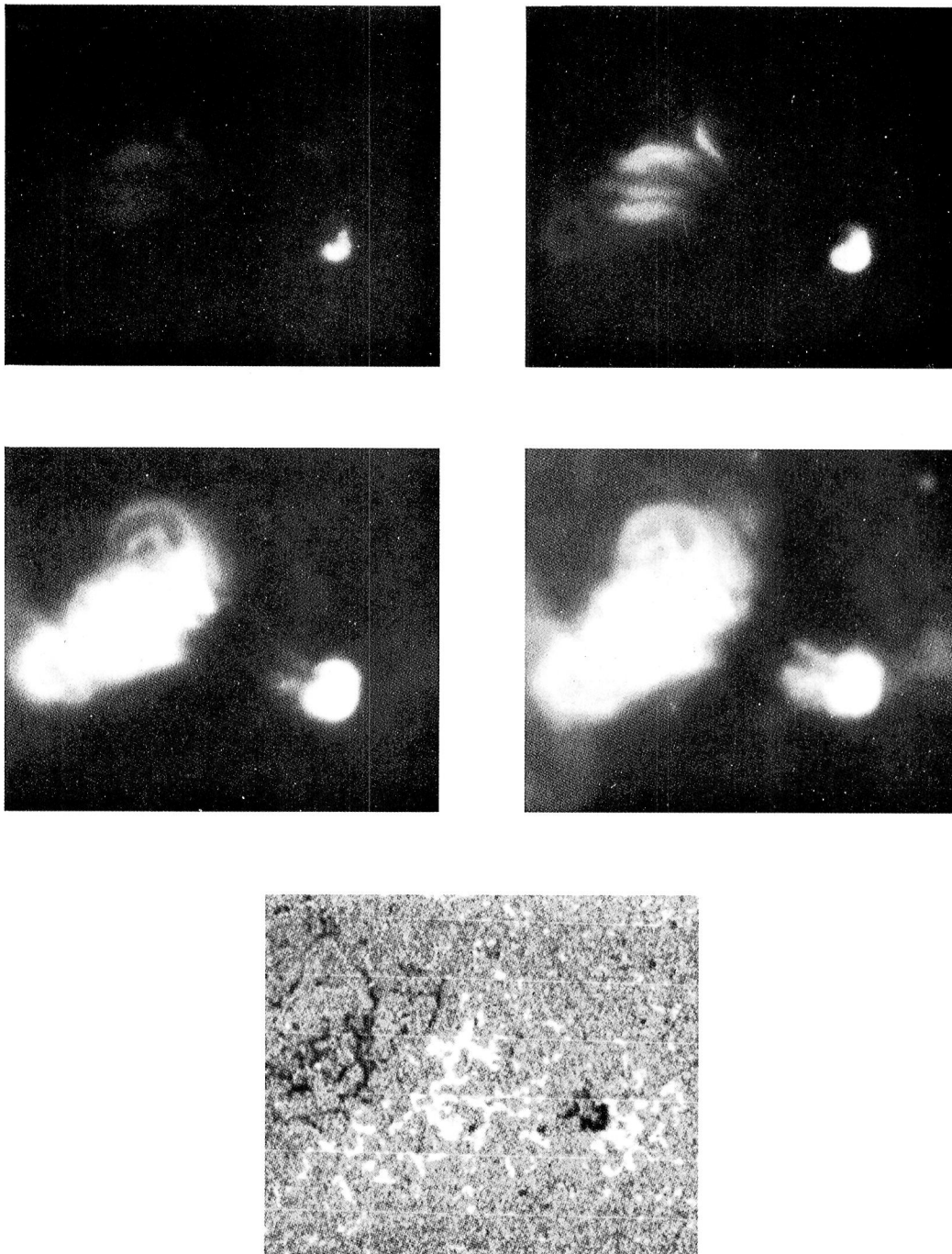
Because magnetic field is divergence-free, field lines cannot begin or end in the photosphere. Thus the emergence of magnetic flux necessarily occurs with the appearance of bipolar magnetic regions (BMRs) at the photospheric level. As much negative polarity flux always reenters the surface at one "footpoint" as it leaves at the other. Initially the two footpoints are very close. As a loop of magnetic flux emerges, the footpoints of the closed structure move steadily apart. Observations between the photosphere and the corona (e.g., in the chromosphere) show linear features in projection between the separating footpoints of a newly born region. These features move upward and dense cool material is often seen draining downward. In general, we visualize the observed emergence process as a twisted subsurface rope of magnetic flux which develops a rising loop or kink that breaks through the surface.

The buoyant rise and expansion of magnetic flux tubes beneath the photosphere has proved a complex and controversial subject. The calculations of Parker (1975, 1977) and Schüssler (1977, 1979) evaluate the rate of rise of isolated flux tubes. Parker (1979a) summarizes these results. The development of flux tubes from diffuse magnetic fields can be explained as instabilities straining an equilibrium state (Tsinganos, 1980, Schmitt and Rosner, 1983, and references discussed therein). Turbulence and nonlinear flows also exclude magnetic flux from the center of convective cells and vortices which fill the convection zone and concentrate it in ropes or strands at the periphery of the cells (Spruit, 1981; Proctor and Weiss, 1982). The release of flux from the base of the convection zone therefore can have both large and small scales. The interaction of this flux with smaller scales of turbulence and convection closer to the surface determines the overall rate and spectrum of flux expulsion from the convective zone. Determining the relative strengths of these effects is important to understanding the drag of magnetic fields on the fluid deep in the convection zone. This is a topic for continuing and future study as closer, clearer views of small flux tubes and emergence events become available.

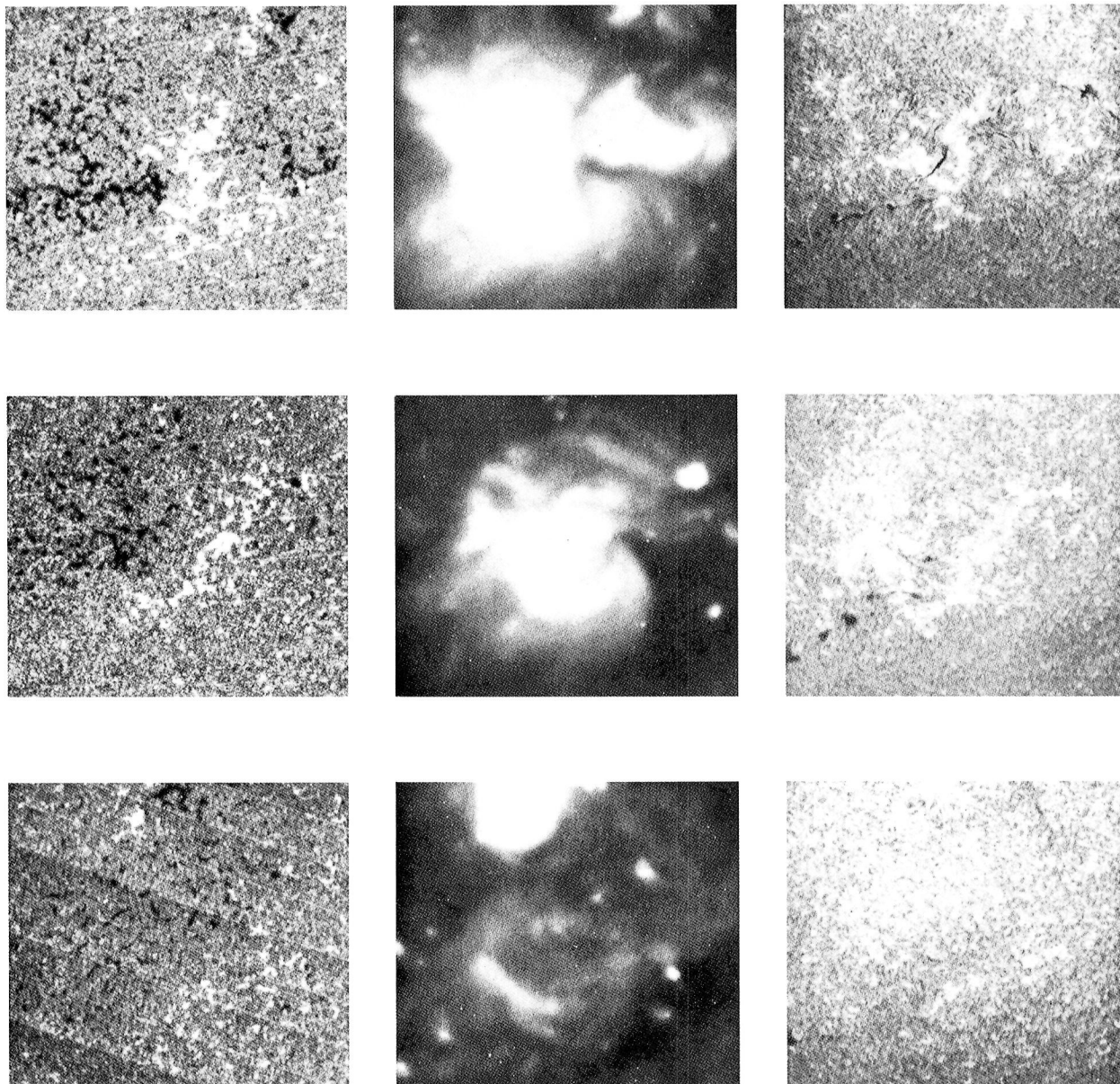
The visible manifestations of flux emergence vary markedly between small and large active regions. The most notable features are the appearance of dark pores (sunspots without penumbrae) at  $\sim 10^{20}$  Mx size and the appearance of increasingly large and complex sunspots within larger regions. At all size scales, magnetic field observations generally show both magnetic polarities emerging in strong-field form close to each other on the surface (also see the section on "Small-Scale Flux Emergence"). Observations of the atmosphere above the solar surface show three-dimensional loop structures filled with X-ray emitting plasma whose footpoints coincide with opposite-polarity magnetic field regions. Figure 3-1 (Rosner et al., 1978) shows four X-ray photographs of a pair of active regions above a magnetogram of the same area. The exposure time increases by a factor of four for each figure from the top left to the bottom right, showing successively larger and fainter loop structures surrounding the bright cores. The younger active region on the right is brighter than the larger and older active region suggesting the importance of transient effects occurring during flux emergence. Figure 3-2 (Rosner et al., 1978) shows three rotations of a long-lived active region, illustrating the coronal dimming as the photospheric magnetic fields diffuse. Each row records three observations taken on subsequent solar rotations. The X-ray emitting plasma loops are anchored in regions of opposite-polarity field as can be seen by overlaying the first and second columns of Figure 3-2. The correspondence between the X-ray visible flux tubes and the chromospheric  $H_{\alpha}$  emission is quite close.

Since the flux emerges in the form of loops or magnetic arches, magnetic tension in the curved lines of force would tend to pull the flux tube downward, even as buoyancy pulls it upward, provided that the horizontal separation of the "anchored" footpoints does not exceed several scale heights (Parker, 1979a). The small-scale effects of this tension are not clearly seen on the Sun, but there are a number of indications that the tension is effective at large scales. Sunspots are seen to rotate across the visible solar surface at a faster rate than the surface plasma and to migrate slowly toward the equator (Gilman and Howard, 1983). These observations are at least consistent with sunspots being connected magnetically to more rapidly rotating, deeper layers. An important question which naturally arises is: How fast can magnetic field be pulled through plasma and what is the drag and heating that result? Quantitative understanding of the solar magnetoconvection clearly depends on these answers.

The speed at which a rope of flux actually emerges is also unknown. If it is slow enough, the quasi-static solutions of Low (1981) should be instructive. If the buoyant flux erupts quickly enough, thin current sheets should form where the emerging flux is pushing less dense plasma



3-1. Four X-ray photographs of a pair of active regions and a magnetogram image of the same area showing opposite polarity magnetic fields in the line of sight as white and black regions. The larger active region is one rotation old and the smaller is four days old in these data taken at 0900 UT on 20 June 1973. The X-ray exposure time increases by a factor of four in each figure from the upper left to the lower right, showing successively larger and fainter loop structures surrounding the bright cores. The energy density of the younger region seems to be an order of magnitude greater than that of the older region.

**MAGNETIC FIELD****X-RAY****H $\alpha$** 

- 3-2. A long-lived active region is shown during three successive solar rotations. The photographs were taken on 2 June (top), 29 June (middle), and 27 August (bottom) 1973. Overlaying the figures shows that the X-ray loops are anchored in regions opposite polarity magnetic field. The sequence of X-ray and H $\alpha$  photographs depicts the response of the coronal plasma to the dispersal of photospheric flux. The X-ray intensities are closely correlated with the intensity of the chromospheric emission.

and flux out of the way. Such configurations should be accompanied by reconnection and plasma heating. If the emergence is even faster, direct shock heating of the chromosphere and corona could possibly result. The rate of emergence is controlled not only by buoyancy but also by the turbulent and macroscopic flows in the fluid which comes up with the flux. If the fluid in the emerging flux tube drains back into the photosphere and chromosphere quickly enough, MHD shocks can result because the Alfvén speed inside the tube will be greater than outside in the corona. Answering these questions clearly has major consequences for coronal heating and other plasma heating and acceleration events.

With respect to larger-scale magnetic flux concentration, it is believed from theoretical considerations that the combination of nonuniform rotation and turbulence produces large-scale ropes of magnetic flux at the very lowest levels of the solar convection zone (Parker, 1975; Gilman, 1983). How then can this presumably large-scale flux emerge through the solar surface in the form of tiny isolated magnetic flux tubes? That magnetic fields in photospheric bipolar regions occur as very small and intense magnetic flux tubes is one of the more puzzling discoveries in solar physics in recent years (Sheeley, 1967; Simon and Noyes, 1971; Howard and Stenflo, 1972). This fine-scale structure appears as flux bundles of a few hundred km in diameter containing  $10^{18}$ - $10^{19}$  Mx of flux, with field strengths approaching 1500 G. Parker (1974a, b, 1978, 1979b) has published a series of papers addressing various mechanisms for the compression of magnetic flux including turbulent pumping, the Bernoulli effect, adiabatic cooling, and concentration over downdrafts.

Turbulence and more organized flows in the upper portion of the convection zone will efficiently shred laminar flux into small-scale ropes or bundles whose cross-sectional dimensions are considerably smaller than the vortices and convective cells which generate them (Galloway et al., 1978; Proctor and Weiss, 1982). Detailed global models of compressible convection, reviewed in Glatzmaier (1983) and references therein, have been developed in several approximations. These models have allowed simulation of the rising "thermals" of hot plasma and magnetic flux and have shown important properties of the fluid dynamics involved. Near the surface in such simulations, however, the local scale height of the fluid has to be unphysically stretched beyond values typical of the Sun. Since the smallest scale heights (~100-200 km) occur just beneath the photosphere on the Sun, the small scales in the magnetic flux which we observe cannot be simulated numerically at the same time as the larger, slower, deeper convective flows because of computer size and speed limitations. Therefore, it is still an open question whether the small-scale magnetic structures are generated preferentially near the surface (i.e. within ~15,000 km) or are the result of slower fluid instabilities originating deeper within the convection zone.

Two well-known hydrodynamic instabilities may play a major role in the relation between flux generation in the solar interior and flux emergence at the solar photosphere: the Rayleigh-Taylor and Kelvin-Helmholtz instabilities. The Rayleigh-Taylor modes could be responsible for the break-up of rather diffuse submerged flux, with magnetic buoyancy subsequently carrying the detached elements of magnetic flux to the solar surface (Parker, 1955, 1975). Both instabilities may have important effects on a rising magnetic flux tube. Unless the tube is strongly twisted, these instabilities, arising from density depletions in the interior of the tube and velocity discontinuities across its boundary, can split the original magnetic flux tube into several magnetic fibers (Tsinganos 1980). This fragmentation should continue until the tube size is reduced enough, or the magnetic tension is strong enough, to resist the destabilizing effects of the hydrodynamic instabilities.



There are many other instabilities which will occur under certain conditions. For example, MHD modes can be unstable on strongly twisted flux tubes and double diffusion, which pits magnetic field diffusion against heat conduction, can also striate a smoothly varying plasma-magnetic field system. A local magnetic flux concentration instability also occurs. Convective instability also can lead to collapse of flux tubes into strong ropes of flux. Peak fields, estimated by balancing magnetic pressure against plasma pressure, are in rough agreement with observation (Spruit, 1981). Computations of magnetic flux concentration by convective eddies and turbulence gives peak fields which are larger than the average background field by a factor of the magnetic Reynold's number (Galloway et al., 1978)

The integrated effect of the small-scale instabilities appears as anomalous transport phenomena, generally represented as greatly enhanced dissipation and diffusion terms in the equations. These, in turn, modify the larger-scale structures in the fluid. The large scales generate flow and field configurations which then become unstable to local modes. This complicated interaction and feedback between the fluid dynamics and local instabilities tends to prevent flux tubes from becoming very highly twisted and unstable gradients from becoming arbitrarily steep.

Assuming that the emerging flux tubes are only weakly twisted, we might account for the absence of large-scale flux elements, since untwisted large flux tubes are vulnerable to the hydrodynamic instabilities. This picture is consistent with the suggestion that large-scale flux elements, such as sunspots, may begin as clusters of very small flux tubes which subsequently merge into larger, single-element flux structures. This merging occurs under the combined action of hydrodynamic effects, such as downward flows (Parker, 1979b), and dissipative phenomena such as resistive tearing.

## UNEXPLAINED OBSERVATIONS

### Small-Scale Flux Emergence

According to conventional wisdom, "all magnetic flux emerges in the form of bipolar loops" (Zirin, 1974). It has been widely accepted that the total flux passing through a region on the surface is constant, unless one end of a flux tube moves into or out of the region, or only one footpoint of a newly forming flux loop appears within the field of view. While this classical view undoubtedly is valid in a majority of cases, recent observations suggest that other, more complex, phenomena also occur. For example, Wilson and Simon (1983) have shown instances in which a small region of positive flux grows and changes its configuration without either the emergence of bipolar loops, the inward movement of detectable flux from outside, or any corresponding emergence of negative flux. In another case, they find that the observed flux associated with a small negative feature fluctuates by ~50% on time scales of 20 minutes, again without observing any lateral addition of flux. Topka and Tarbell (1983) report transients in emerging loop events, while Patterson and Zirin (1981) and Mayfield and Chapman (1981) describe several seen during flares. Sheeley (1982a) has reported flux changes in Kitt Peak and San Fernando magnetograms, while Bumba (1981) has studied the initial stages of local field formation, and finds it very difficult to explain his data solely in terms of passively emerging flux tubes.

One possible explanation for these apparently unipolar phenomena might be the existence of many small bundles of both positive and negative flux below the resolution limits of the magnetograms. In this case, a positive polarity region could grow by the addition of unresolved positive bundles at the periphery of the observable region. The implied excess of negative



bundles could transmit the excess negative flux away from the growing region of positive polarity undetected under the right conditions of flow or reconnection. Wilson and Simon (1983) observe a significant correlation between magnetic field gradients and gradients in the velocity field, which suggests that a velocity shear between regions of positive and negative flux may be directly involved in the observed magnetic transients.

Recent observations (Topka and Tarbell, 1983) have yielded high resolution composite profiles including magnetic fields in the line of sight and luminosity variations which show up the granulation. A large number of flux bundles are seen in the intergranular lanes, but the central regions of the granulation cells are not always devoid of flux. Figure 3-3 is a set of Doppler and magnetic images (color) overlaying a set of line-center and continuum images (black and white), all taken in the Fe I 5250 line, which show the convection patterns of the granulation, the concentration of magnetic elements in the intergranular lanes, and evidence for a small-scale flux emergence event. A better understanding of these new data is now being sought (Wilson, 1983; Wu et al, 1983).

#### Anticorrelation of Small- and Large-Scale Emergence

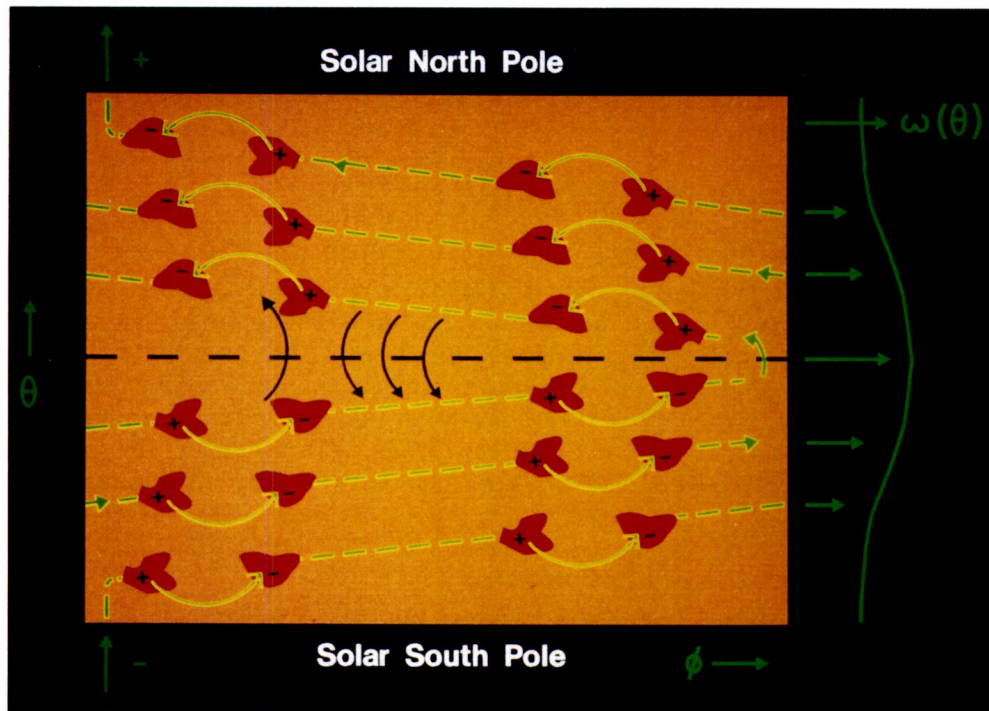
From rocket flight data obtained during the period 1970-78, Golub et al. (1979) deduced a strong anticorrelation between the sunspot number and the number of X-ray bright points which are felt to mark small-scale magnetic flux emergence events. This result was confirmed by additional recent rocket X-ray observations (Davis, 1983), but it should be noted that Martin and Harvey (1979) find just the opposite result using magnetograph data - small region flux emerging in phase with the cycle.

Such an anticorrelation between large- and small-scale emergence of magnetic flux would mean that the total flux emergence rate throughout the activity cycle is nearly constant. One could then characterize the cycle as an oscillation in the wavenumber distribution of the emerging flux, rather than in the quantity of flux produced. The resulting picture would be consistent with identifying bright points with the emergence of turbulently regenerated flux left over from the large complexes of activity which populate the sunspot belts during the peak activity periods. In the low activity portions of the solar cycle, large volumes in the Sun's convective zone would then necessarily exist where fields of opposite polarity are intermingled and continually reconnecting without large magnetic structures causing big active regions.

On the other hand, if the X-ray bright points do appear in phase with the sunspot cycle, it means small-scale flux regeneration in the absence of larger-scale background flux elements can not be as important an effect.

#### Organization in Large-Scale Emergence

For many years solar physicists have debated the existence of "active" longitudes rotating with the Sun on which there is a greater than average likelihood that a bipolar magnetic region will appear. Although the existence of one or more favored longitudes is still open to question, it has become clear that active regions often are organized into very large patterns called complexes of activity (Bumba and Howard, 1965; Gaizauskas et al., 1983). Within a complex of activity, which typically has a diameter of 300 Mm, the individual regions form successively in a temporal sequence, spreading out from a central longitude. The resulting complex of activity occupies a single extended longitude zone for a year or more.



3-3. Four composite images of an emerging flux region in the photosphere, taken in the core, wings, and nearby continuum of the line Fe I 5250. The tickmark spacing is 2 arcseconds (1440 km). Shown in gray tones are the line center (upper left and right) and continuum (lower left and right) images. Overlaying these images are Doppler (upper and lower left) and magnetic (upper and lower right) maps, with the line-of-sight velocity and magnetic field toward the Earth shown in blue and away from the Earth shown in red. Several interesting features are evident:

- The velocity-continuum image (lower left) shows blue appearing over brighter regions, indicating upward gas flow in the center of granules, while the dark granule boundaries often exhibit a downward gas flow (red). Note also the large-scale downflow regions, which are cospatial with the large-scale magnetic field structures seen in the magnetic-continuum image (lower right).
- The magnetic-continuum image (lower right) shows that whenever the magnetic elements are small, they occur in the dark intergranular lanes. Large concentrations of flux, on the other hand, cover the granulation pattern.
- The magnetic-line-center image (upper right) shows a small dark region, marked by the intersection point of the arrows, which appears in the line-center images but not in the continuum images. The magnetic images show that the line-of-sight field is negligible in the dark region, but that field is present and oriented out of the surface (blue) to the left of the dark region and into the surface (red) to the right. The velocity images show that there is a strong upflow (blue) in the dark region, while downflows (red) are found in both neighboring magnetic areas.

The interpretation of these observations is that a rising magnetic loop connects the small opposite polarity regions. The dark region is dense material being carried up from below the height of the continuum, and the material is draining down along the magnetic field. Within twenty minutes of these exposures, the dark feature had disappeared. (Figure provided by A.M. Title)

As the result of such organized complexes, one sees a number of alternating bands of weak positive and negative polarity fields, stretching over the circumference of the Sun. Perhaps 6-10 of these features, interrupted here and there by active regions, extend around the circumference of the Sun at active region latitudes. The boundaries of these features, although somewhat distorted by the shear of differential rotation, do not show nearly the effects of the differential rotation that one could expect over the pattern lifetimes which often exceed 10 rotations. This rigid body rotation of magnetic patterns (but not individual features) has been recognized for a long time (Bumba and Howard, 1965), and is often seen in the motions of other features, such as coronal holes and the interplanetary magnetic field.

These observations strongly suggest a global underlying pattern in the way magnetic flux comes to the surface, either a signature of an underlying cellular convective pattern or the imprint of a global scale instability, which starts the flux rising. The pattern is dominated by a narrow band of modes with generally eight or nine wavelengths around the circumference of the Sun. The recent observations also highlight the differences between the motions of magnetic field patterns and of the underlying fluid, as discussed in "Models for Flux Emergence" above. Determining this flow accurately and understanding the physical mechanisms which give rise to it is clearly a major challenge to both observation and theory and the answer will be a step toward understanding how the solar dynamo works.

### III. DYNAMICS OF SURFACE MAGNETIC FLUX

Because of the observed intermittency of emerging flux on all scales and the local intensity of the flux bundles, the discussion of flux emergence in Section II naturally concentrated on small-scale structures and the third focal question, what is seen at the limits of resolution? When one studies the dynamics of magnetic flux which can be followed on the solar surface for any appreciable time (say greater than 1 week), consideration of small bundles and granule-sized ephemeral regions has to give way to consideration of large-scale patterns. The problem is then the separation of weak average behavior from the noisy background of the strong, small-scale transients.

#### MAGNETIC FLUX TRANSPORT

##### Physical Processes Which Transport Flux

The motions of magnetic fields on the Sun and the motions of the solar plasma are strongly coupled through the high electrical conductivity of the atmosphere. As a consequence of this coupling, the magnetic flux is moved bodily by large-scale coherent fluid flows and is structured by smaller-scale motions and convective turbulence in the convection zone of the Sun, as described briefly in "Models for Flux Emergence" above. The large-scale concentration and dispersal of the flux threading the surface of the Sun results from the interplay of differential rotation, meridional circulation and turbulent convection.

The rotation of the Sun has been studied extensively (see Howard, 1978, and references therein). Differential rotation, the variation of the surface rotation rate with latitude, is clearly observed in the motions of the photospheric gas and surface tracers such as sunspots. Differential rotation shears the extended active regions and enhances the diffusive mixing of opposite-polarity

fluxes by lengthening the magnetic neutral line between them. The variation of the rotation rate with depth in the atmosphere has been inferred from the observation that magnetic features and the surface plasma do not rotate at the same rate. This variation is an important parameter in models of the solar magnetic dynamo (Stix, 1981), and has recently been inferred from observations of global acoustic oscillations of the solar atmosphere (Hill et al., 1983).

Modulations of the rotation rate of the Sun on long time scales have been reported recently. Eddy et al., (1976) have measured a lower rate for the thirty years preceding the Maunder Minimum (1645-1715) than in this century. The early data used in this study are poor by modern standards, and this result should be considered uncertain (Abarbanell and Wöhl, 1981). On shorter timescales periodic modulations are seen in the solar rotation rate. Changes related to the solar cycle have been seen in Doppler velocity measures (Howard, 1976; Livingston and Duvall, 1979) and in sunspot tracer measures (Gilman and Howard, 1983). The rates increase near solar maximum and minimum with amplitudes of variation of about 1% for spots and 2% for the photosphere. The sunspots rotate faster than the photosphere by about 4% on average.

The Sun also appears to exhibit torsional oscillations with a period of 11 years, amplitudes of about 10 m/s, and with 1 or 2 zones per hemisphere (LaBonte and Howard, 1982a). The oscillations are symmetric about the equator, suggesting that they are the surface manifestation of some global convection pattern. There is evidence for connection between the shorter wavelength torsional mode and magnetic activity: the activity zone is centered on the latitude of the torsional shear. The amplitudes of the torsional modes seem to be independent of the level of magnetic activity which suggests that they are hydrodynamic in origin, and not dependent on the magnetic fields generating the dynamo.

Meridional circulation affects the concentration and dispersal of magnetic flux by the latitude dependence of its direction and magnitude. The direct detection of meridional flows is difficult; the velocities involved are near the limit of resolution of the Doppler measurements, and the meridional flow signatures are not clearly distinct from limbshift effects, radial inflows in active regions, and other large-scale velocity fields (LaBonte and Howard, 1982b). Evidence is accumulating for a poleward meridional flow of magnitude 10-20 m/s (Beckers, 1979; Duvall, 1979; Howard, 1979; Beckers and Taylor, 1980; LaBonte and Howard, 1982b), although detection of an equatorward flow has also been reported (Perez Garde et al., 1981).

The existence of a poleward flow has also been inferred from motions of magnetic features. Large-scale, directed flows toward the solar poles have been observed in weak surface flux (Howard and LaBonte, 1981) and also in large-scale  $H_{\alpha}$  patterns (McIntosh, 1979), and an analysis of the motions of polar filaments suggests a poleward flow of about 10 m/s (Topka et al., 1982). Measurements of sunspot motions, on the other hand, have shown small amplitude equatorward flows (Gilman and Howard, 1983). Presumably this implies a return motion at the depth where the spot fields are anchored. These results are all consistent with a circulation pattern of upflows at equatorial latitudes, downflows at polar latitudes, and poleward flows at the surface and equatorward return flows at depth at intermediate latitudes.

Convective turbulence on the Sun appears to be structured on at least four distinct spatial scales associated with corresponding distinct depths in the solar atmosphere (Simon and Weiss, 1968): giant convection cells in the lower convection zone (now called giant granulation), supergranulation and mesogranulation (November et al., 1981) in the upper convection zone, and granulation in the photosphere. The large-scale complexes of activity and their attendant surface

velocity features (section on "Organization in Large-Scale Emergence") may be the surface manifestations of giant cells deep in the convection zone. The concentration and turbulent dispersal of the surface magnetic flux is due to the shifting circulation patterns of the supergranulation, mesogranulation and granulation convection cells.

Turbulence theories have been applied to help explain the growth, structuring, and dissipation of magnetic fields on the Sun. Knobloch and Rosner (1981) discuss MHD turbulence as applied to flux transport and attempt to relate theory to observation. Their conclusions are:

- the motions must be three-dimensional,
- the effective diffusivities are much larger than molecular diffusivity, and
- the dynamical processes which control the magnetic field spectrum occur below  $\sim 15,000$  km.

While turbulence appears as an anomalously large diffusive transport term when viewed macroscopically, it is clear that the intense, small-scale "salt and pepper" structures which are seen could not exist if the extrapolation of the observed large-scale turbulent diffusivity were carried out blindly to these smaller length scales. Indeed, the effective diffusivity at short scales may not even be positive (Kraichnan, 1976; Knobloch, 1977).

What remain to be determined are the structure and dynamics of the giant and mesogranulation convective patterns which can be assumed to play an important role in driving the 15,000 km scales. Is the large-scale convection regular and steady, the solar cycle being just a weak, varying magnetic perturbation which rides piggy-back on a nearly steady convection, or is the convection pattern itself changing measurably with an 11-year cycle? The shorter scale convections, granulation, mesogranulation and supergranulation, are clearly transient but there is no reason why the larger scale patterns may not be permanent. The 11-year period of the torsional oscillations, if they are indeed hydrodynamic in origin, argues that the dynamo is being driven with an 11-year period rather than self-consistently adopting that period as a result of magnetic forces interacting with the Sun's global convection.

### A Model for Flux Transport

The nonstationary character of the global supergranulation and granulation patterns gives rise to a random walk of magnetic flux over the surface of the Sun (Leighton, 1964). Leighton incorporated this process into the first quantitative model for magnetic flux transport in the photosphere. He assumed a continuous distribution of magnetic lines of force over the Sun and considered the evolution of this distribution subject to differential rotation and diffusion. From an assumed rate of emergence of sunspot flux as a function of latitude over the solar cycle, Leighton evaluated the diffusion constant in his model by fitting the calculated temporal variations of the strength and boundary location of the polar fields to those inferred from observations of polar faculae and polar prominences. He found that a value of  $770\text{--}1540 \text{ km}^2/\text{s}$  fit the observations rather well.

Some years later, Mosher (1977) obtained a significantly lower estimate of  $200\text{--}400 \text{ km}^2/\text{s}$  for the diffusion constant from measurements of the rate of areal spreading of individual evolving active regions. He suggested that a poleward meridional flow of  $2\text{--}5 \text{ m/s}$  would assist the diffusion sufficiently to account for the observed variations in the polar field strengths and boundary locations.

Mosher's conclusions and the accumulating direct evidence for a meridional flow motivated recent efforts by Sheeley et al. (1983) to generalize Leighton's transport model to include the effects of meridional surface currents, and to apply the model to a detailed study of the evolution of individual magnetic active regions on the Sun. These investigations mark the first application of a flux transport model to a fully quantitative study of the Sun's large-scale magnetic fields using solar magnetic data as input. The most recent results of these studies are consistent with a poleward flow of 10-20 m/s and a diffusion constant of about 400 km<sup>2</sup>/s (Boris et. al., 1983), substantially in agreement with the work of Mosher.

The characteristics of the Sun's polar magnetic fields are determined to a significant degree by the flux transport process. When a bipolar magnetic region emerges in the sunspot belt, the footpoint of the same polarity as the corresponding polar field almost always emerges in front of the opposite polarity footpoint. The flux of opposite, or "following" polarity then rotates behind the "leading" active region on the Sun and generally emerges with its centroid a little further from the equator. The diffusion and convection of active region flux of predominantly "following" polarity toward the polar regions through the sunspot cycle leads to the cancellation and disappearance of the old polar flux and its replacement by new flux of opposite polarity.

For a meridional flow speed and turbulent diffusivity of the order of 10 m/s and 300 km<sup>2</sup>/s, respectively, the convection time is an order of magnitude smaller than the diffusion time. As a result, the meridional flow sweeps up the magnetic flux and concentrates it at the poles until the field gradient is sufficiently large for the diffusion and convection to come into balance. The details of the concentration and strength of the fields are sensitive to the meridional flow profile in the polar regions, and both of these quantities are difficult to determine observationally. An outstanding issue is to reconcile the concentration of polar fields predicted by Svalgaard et al. (1978), who extrapolated observations from lower latitudes to obtain an estimate of the polar field strength of about 10 G, and direct measurements of field strengths of less than 5 G (Howard and LaBonte, 1981).

These data raise an important theoretical question: How are we to interpret these large-scale flux migration observations in terms of the dynamo and the solar cycle? How does the flux which leaves the photosphere in the following active regions, which on average migrates northward, connect up across the equator and deep within the convection zone with the flux from following regions in the Southern Hemisphere? Thus magnetic reconnection from resistive tearing seems necessary to provide deep reversed flux which initiates the second half of each solar cycle. An alternative to subsurface reconnection is to have all the flux necessary to reverse the average dipole come in from infinity - as it would if the magnetization of a conducting sphere were being reversed by slowly changing the toroidal (equatorial) current.

## MAGNETIC FLUX STRUCTURE

### Flux Concentration

Soon after the discovery of solar supergranulation (Leighton et al., 1962) it was recognized (Parker, 1963; Simon and Leighton, 1964) that these cellular convective motions organize the surface magnetic fields into a network pattern (Leighton, 1959) with corresponding brightness enhancements which we call the chromospheric and photospheric networks (Deslandres, 1899; Chapman and Sheeley, 1968). Originally it was thought that these flow patterns could concentrate the magnetic field up to the limit imposed by equipartition of kinetic and magnetic



energy densities ( $B^2/8\pi = 1/2 \rho v^2$ ), about 20 G at the photospheric level. Detailed observations (Sheeley, 1966; Beckers and Schröter, 1968a, b; Ramsey et al., 1977), however, raised the observed intensities successively from 500 G to 1000 G and now 1500 G. In addition, models suggest that such fields may reach 2000 G (Stenflo, 1976). Whereas early magnetic maps of the Sun indicated a weak, relatively uniform solar field of only 1-2 G outside sunspots and pores (Babcock and Babcock, 1955), recent work suggests that most of the solar flux occurs in small concentrations of very strong fields (1-2 kG). The most recent and exciting observation (Topka and Tarbell, 1983) confirms a long-held expectation that even granules, the smallest of the convective structures seen on the Sun, should be able to concentrate flux at their boundaries (Figure 3-3).

If we assume pressure balance between undisturbed nonmagnetic photospheric gas and an imbedded empty flux tube we require a field strength at the solar surface of about 1500 G for a tube supported by magnetic pressure only. Spruit (1981) and Parker (1974a, b, 1978, 1979b) consider the concentration of magnetic flux into flux tubes by a number of mechanisms including adiabatic cooling, turbulent pumping, and concentration over downdrafts. These analyses are consistent with the present observations of strong, concentrated magnetic fields.

It is relatively easy to model the motions of randomly placed flux tubes from below the photosphere up to the surface, and then to the boundaries of supergranule cells (Meyer et al., 1979). It is more difficult to describe the complex processes which amplify and concentrate the field strengths to their limiting values (Galloway and Weiss, 1981; Proctor and Weiss, 1982; Spruit, 1981). Detailed model calculations successfully predicted the newly observed magnetic flux concentrations in the intergranular lanes separating granules (Topka and Tarbell, 1983). Indeed, the concentration of flux by fluid dynamic effects in and near the photosphere is indirect support for the notion that the observed active region fields result from convection and turbulence on a larger scale deeper in Sun.

### Weak Background Fields

In the first years of magnetograph observations it was assumed that the extended weak fields that were observed actually existed at the strengths measured on the Sun. Indeed, the magnetograph was developed originally to measure the "general field" of the Sun - a global dipolar field, analogous to that of the Earth, which astronomers for many years assumed to exist on the Sun.

As the spatial resolution of magnetograms improved it became clear that magnetic fields are confined mainly to tiny bundles of very strong fields, as described above. Even the fields measured in the polar regions of the Sun are made up of such strong fields. This discovery solved a theoretical problem that had worried solar physicists for many years: how could the weak (1 or 2 G) fields that were observed not be greatly amplified by the strong convective motions of the granulation pattern?

Clearly there are weak spatially averaged solar magnetic fields because they have been measured for years. The physical question that arises is then: are there any actual regions of weak field on the Sun at all? There are no observations which show directly that very weak fields are present on the Sun. In an analysis of magnetic field measures in two spectral lines having different excitation properties, Howard and Stenflo (1972) found an upper limit of 10% as the fraction of magnetic flux that could exist as weak fields. This means that 1-2 kG fields over 1% of the solar surface would carry at least 10 times as much flux as is contained in the average 1-2 G solar field. This picture is fully consistent with flux being swept up and compacted into bundles at the granulation and supergranulation cell boundaries.

Nevertheless, it seems unlikely that fields are exactly zero between the bundles. In the intergranular lanes and the supergranular network the fluid dynamic convergence effects are all finite so flux would not be completely removed from one region any more than it is infinitely compressed in another. This becomes important because the apparent coherent migration of flux at substantial speeds through the photospheric flow is possible when a nearly horizontal weak average field is present. Consider a large flux tube which is inclined at only a slight angle to the horizontal as it rises convectively at a modest rate. If the angle is shallow enough, the surface intersection of the flux tube could appear to move an order of magnitude faster along the surface.

Even if the weak horizontal flux tube is fragmented into a number of intense flux tubes along with all the strong vertical flux, reconnection could easily ensure that the small excess flux of one polarity would move freely about (Sheeley et al., 1975). In flux transport models (Leighton, 1964; Sheeley et al., 1983) the correctly phased sequence of small sources would properly represent this coherent emergence phenomenon. Simple extensions of this phenomenon could account for unipolar regions apparently forming or dispersing quickly and sunspots and active regions moving relative to the surrounding fluid and yet apparently leaving no wake.

#### IV. DISAPPEARANCE OF MAGNETIC FLUX

Without a mechanism to remove magnetic flux from the photosphere, the continual eruption of new flux would eventually lead to very strong magnetic fields everywhere on the Sun. This is not observed. Not only does the eruption continue more or less uninterrupted year after year and cycle after cycle, but the large-scale field also remains quite weak. For example,  $10^3$  bipolar magnetic regions with an average flux of  $5 \times 10^{21}$  Mx would produce  $5 \times 10^{24}$  Mx of each polarity. Spread over a solar hemisphere, this would give a mean field strength of 200 G, compared to the 1 G field that is observed. Evidently, the Sun finds an efficient way of getting rid of photospheric flux.

Indeed, magnetograph observations show that a substantial fraction of the emerging flux (perhaps 10% or more of the flux that erupts in active regions) somehow disappears on a time scale of only 1 day (Howard and LaBonte, 1981). Sheeley et al. (1975) show a number of instances of how this happens by comparing magnetograms and Fe XV spectroheliograms of new active regions on successive days. They identify reconnections of flux eating away at newly erupted bipolar magnetic regions. It is important to recognize that this flux disappears approximately where it erupts, not having had sufficient time for it to spread to more remote parts of the Sun. Thus, on the average, only a small fraction of the flux from each active region is available to contribute to the production of the large-scale surface field, and the overwhelming majority of the emerging flux disappears in the sunspot belts shortly after it appears.

#### THEORETICAL CONSIDERATIONS

Although fresh magnetic flux emerges in large quantities as the result of the solar dynamo, some unknown combination of mechanisms removes or cancels the emerged flux efficiently. Some kinematic considerations of magnetic field topology will help to clarify the issues presented by pic-



ture. Because the magnetic field is divergence-free, photospheric flux can disappear only by merging field lines of opposed polarity. In this way, the total fluxes of both signs are reduced simultaneously and in equal amounts. It is important to stress that the actual reduction of magnetic flux passing through the photosphere occurs as a mutual annihilation of opposite polarity regions.

We consider three mechanisms for the removal of magnetic flux from the photosphere:

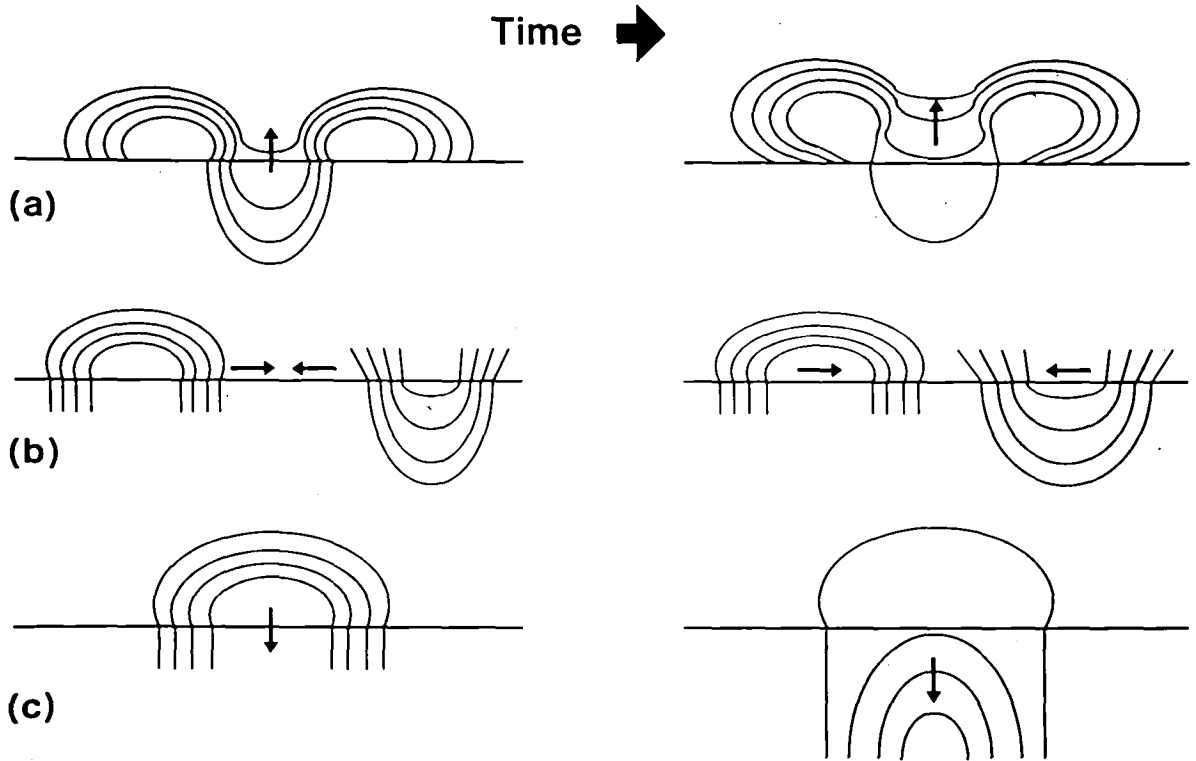
- (1) Release of a partially submerged flux loop into the corona,
- (2) In situ dissipation of flux and the current which causes it, and
- (3) Submergence of a partially exposed flux loop in the photosphere.

Mechanisms (1) and (3) can operate in a medium of infinite electrical conductivity, while mechanism (2) requires finite resistivity. The release of a partly submerged loop into the corona is just flux disappearance through emergence. The release may occur for a closed configuration, i.e., a magnetic bubble, or for the submerged part of a multiply re-entrant field line (Figure 3-4a). In situ dissipation of magnetic flux through resistive diffusion, perhaps in conjunction with dynamical reconnection, may take place for either partly exposed or partly submerged field lines (Figure 3-4b). The submergence of partly exposed field lines is sketched in Figure 3-4c.

What are the signatures of these processes? The rise of an inverted loop would accompany both the release and the dissipation of partly submerged field lines, while the collapse of a bipolar loop would accompany both the dissipation and the submergence of partly exposed field lines. Distinction between the geometries, partly exposed versus partly submerged, is potentially possible with high-resolution vector magnetograms. Distinguishing the dissipative and non-dissipative processes would require simultaneous velocity and magnetic measurements, to determine whether the convective loss of flux is sufficient to account for the observed reduction of flux in the region of activity.

The consequences of the three mechanisms are quite distinct. Recent estimates (Book, 1981) indicate that the release of flux into the upper atmosphere could provide sufficient flux, mass, and energy to heat the corona and power the solar wind. In situ dissipation, almost certainly in the form of reconnection in the network, would heat the photosphere and liberate much less flux and energy into the corona. Flux submergence would have important implications as a significant element in the removal of surface flux; much of the "new" flux emerging on the Sun would then be just recycled "old" flux, lowering our estimate of the rate of production of new flux by the solar dynamo.

Are these mechanisms all likely to be effective on the Sun? Flux release is a natural consequence of buoyant emergence and reconnection and seems to be the best candidate for flux disappearance. The high electrical conductivity of the photosphere requires that the magnetic field vary over very small spatial scales in order to dissipate flux on a time scale of days. Classical resistive dissipation thus seems unable to provide enough flux disappearance unless natural fluid motions automatically generate the small scales required. Since, however, the surface magnetic flux does exist in intensely concentrated bundles, the stringent conditions for rapid resistive tearing and annihilation may be met. Finally, the combination of fluid downflows and magnetic tension must overcome the strong buoyancy of the emerged loop in order for flux submergence to take place.



### Magnetic Flux Disappearance

3-4. Mechanisms of flux disappearance from the photosphere. Figure 3-4a depicts a loop of submerged flux connecting two exposed coronal loops. Magnetic tension snaps the flux upward as the denser, heavier photospheric gas is left behind. Figure 3-4b shows the annihilation of flux at the surface as an o-point disappears. Figure 3-4c shows the resubmergence of a partly exposed loop of flux. The signature of all of these modes of flux disappearance is two footpoints of opposite polarity coming together and annihilating.

Disappearance of flux from the surface of the Sun is a largely unexplored problem. The observations are difficult because of the high spatial and temporal resolution required, and quantitative estimates of the rate of flux disappearance, from theoretical considerations, apparently do not exist. A better understanding of these processes is essential to our understanding of the solar cycle.

## OBSERVATIONS OF FLUX DISAPPEARANCE

### Small-Scale Disappearance of Bipolar Loops

During the Skylab mission in 1973-74, simultaneous coronal and magnetic observations of several average- to small-sized active regions showed similar patterns of evolution. In each case, a substantial fraction of the active region flux appeared to diffuse across the neutral line and be annihilated, while the centroids of the remaining flux of each polarity drifted progressively farther apart. One would expect this annihilation to be accompanied by a continual collapse of closed field lines moving downward from their original positions arching across the neutral line (cf. Sheeley, (1982b)), as discussed above. Indeed, as the photospheric flux disappeared over a period of several days, the corresponding Fe XV coronal loop system steadily contracted in size until neither the loop system nor the photospheric magnetic region remained on the final day.

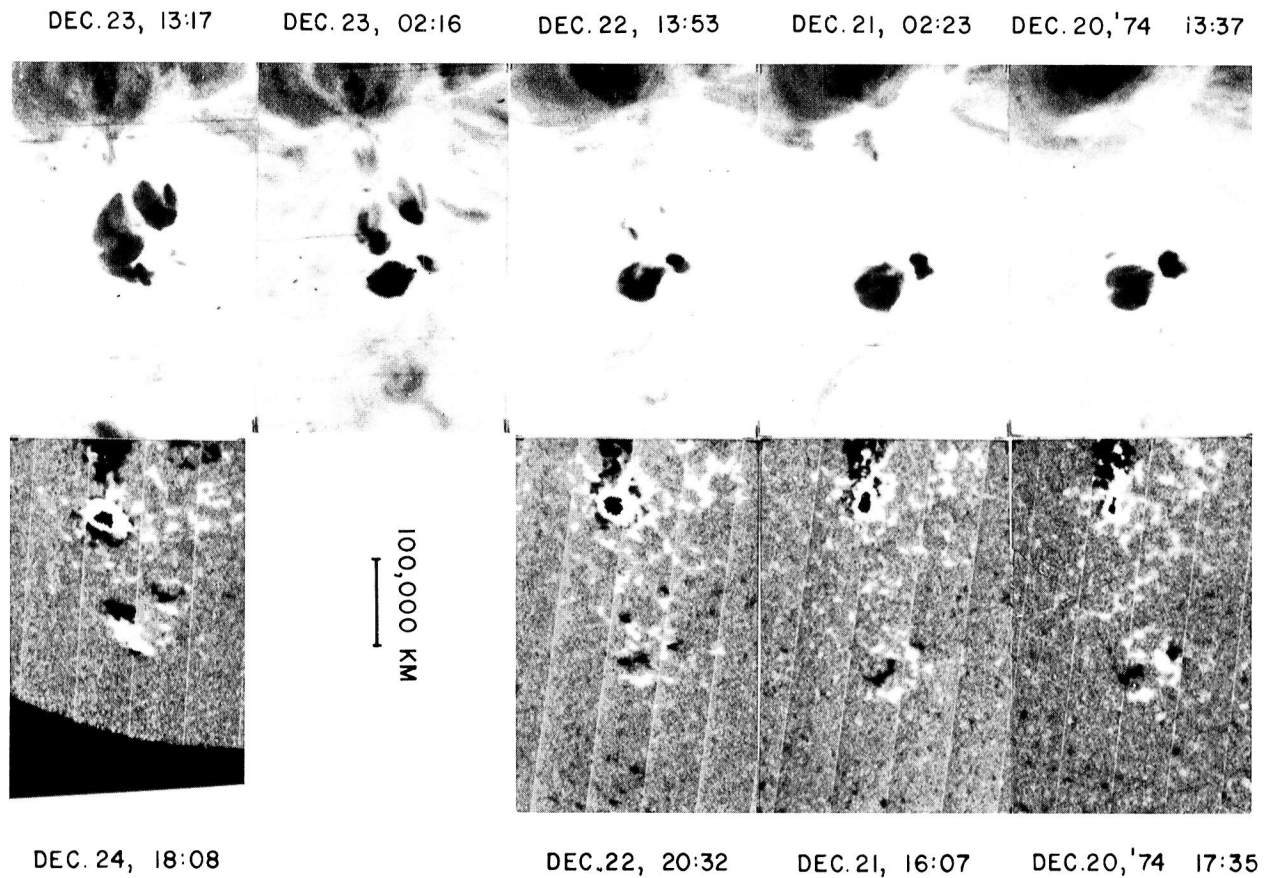
Figure 3-5 shows examples of both reconnection and magnetic flux disappearance observed during the interval 20-24 December 1973 (Sheeley et al., 1975). A pair of new bipolar magnetic regions first appears on 22 December in a region of predominantly positive (white) flux separating a large sunspot from two existing bipolar magnetic regions. Reconnections to the positive background flux begin on December 22nd and become more pronounced on the 23rd and 24th. As each of the two new little regions appears on the 22nd, the field lines and consequently the Fe XV emissions are very localized. Reconnections of the plus and minus flux from each footpoint to different external plus and minus flux tubes progressively broadens the area of coronal emissions as shown in the two frames in the upper left of Figure 3-5.

The northernmost of the two originally existing bipolar magnetic regions decays and is essentially gone by Christmas Eve. In this case the disappearance of the small active region occurred in the presence of the even larger sunspot and other small active regions. In particular, the small region erupted in a large positive polarity region. As a result the negative polarity part of the small region was annihilated in only a few days. Its corresponding positive polarity flux is presumably still present, but one can not track it because it is indistinguishable from the positive flux of the larger region in which it was embedded.

### Sunspot Erosion

The decay of sunspot magnetic flux seems to represent a special case of magnetic flux disappearance (Wallenhorst and Howard, 1982; Wallenhorst and Topka, 1982). Spots disappear rapidly without disturbing nearby weaker plage magnetic fields. The flux of sunspots does not seem to be converted to or get mixed with plage flux; the sunspots seem to disappear in place.

A plausible model for such flux disappearance can be imagined in terms of small-scale, strong-field elements transported by local convective motions or successive reconnections to the vicinity of the spot and cancelling a small amount of its flux. However, this does not fit the generally accepted picture of the distribution of fields on the solar surface, and it suggests that



3-5. A time-lapse sequence of Fe XV coronal spectroheliograms and photospheric magnetograms illustrating both the reconnection of emerging bipolar flux and the disappearance of other bipolar flux in the vicinity. Notice that the smaller of the existing bipolar magnetic regions (at 2 o'clock from the major active region in the lower center of the pictures) has virtually disappeared by 24 December 1973.

sunspot fields are somehow different from the plage fields in which they are imbedded. It may be that these plage fields, which are hypothesized to cancel the spot fields, are related either to the intergranular fields detected recently (Tarbell et al., 1983) or to the fields seen in the centers of supergranules (Livingston and Harvey, 1971).

The explanation for this behavior may lie in the lack of adequate observational resolution. Could there be a sufficient amount of opposite polarity magnetic flux in unresolved small-scale elements in the vicinity of the spot to annihilate the enormous spot flux? If the answer is yes, a magnetogram of an outstanding spot with coherent flux of a given polarity (Figure 3-6) ceases to be impressive; for the same magnetogram also carries a large amount of flux of the opposite polarity in small unresolved structures. An alternative hypothesis is that abundant unresolved elements of the opposite polarity are not present, but that the spot flux itself dissolves into unresolved flux elements that eventually move out of the observed region, and are annihilated by opposite polarity fluxes elsewhere. This explanation in its simplest form seems unlikely since the neighboring structures of a decaying spot appear to remain unchanged during this evolution.

Sunspot erosion is an outstanding issue in solar magnetic activity. The decay of sunspot flux is sometimes observed as visible flux elements leaving the spot. It is crucial to improve resolution to see if the postulated unresolved flux elements exist. At the same time, the theoretical question of limits on the size of such elements can be addressed, based on our current understanding of convection in the photosphere.

## V. SUMMARY

To understand the evolution of solar magnetic flux we must answer the fundamental question: How does the solar dynamo work? Recent observations certainly influence models of the dynamo but we have not been able to deduce the subsurface flows which control it. A reasonably complete understanding of the dynamo will only evolve from synthesizing clues uncovered in research aimed at the four issues identified in the Introduction:

- (1) How does flux emerge on the Sun?
- (2) How does magnetic flux disappear?
- (3) What is seen at the limits of resolution?
- (4) What large-scale patterns are detected?

Numerous aspects of the first two questions arose repeatedly in theoretical attempts to describe what occurs during the life cycle of magnetic flux on the Sun. The last two questions are observational and arise from needed advances in our ability to resolve and measure small time- and space-scale structures and in our ability to extract weak, large-scale patterns from noisy background over a long period.

Progress on the observational side has been rapid but we still do not have enough detailed, simultaneously obtained information about small-scale structures and dynamics to distinguish among many possible interactions and configurations. Specifically, it is crucial to see into the subarcsecond regime (50-500 km) and to identify and record phenomena in seconds rather than minutes to understand the dynamics of emergence and disappearance of flux. Small structures (~100 km) and correspondingly short timescales (~10 s) in the solar plasma temperature, density, velocity, and vector magnetic fields seem to control much of the dynamics that we do see.

Theory and numerical modeling are being used to bridge this gap, making increasing use of supercomputers and simulation techniques. Macroscopic observables are estimated computationally assuming certain mechanisms and configurations which have been postulated but which have not yet been verified by direct observation. This approach will bear more fruit as corresponding observational improvements are made.

Clearly one direction of future research will be to obtain adequate resolution in time and space of the elemental flux tubes as they emerge. Rare pictures under conditions of particularly good viewing yield marginal resolution of some 100-200 km structures. To study emergence, submergence, and reconnection quantitatively will require not only better spatial resolution but also the capability to obtain a whole time series of high resolution shots reliably. Space-based telescopes and instrumentation meet some of these criteria in a way that ground-based observation cannot match.

While theory can lead observation in some areas, it necessarily lags behind in others. Our theoretical understanding does not even encompass some large-scale solar phenomena which have become well established in the last two or three decades. Magnetic activity and related transient phenomena seem to present particularly difficult theoretical problems. Examples of observations waiting for a quantitative theory are coronal heating, solar prominences, coronal transients, the equatorward migration of the sunspot belts from the beginning to the end of each solar cycle, the acceleration of the solar wind, and the solar dynamo itself.

The problem of magnetic flux emergence is receiving more attention with the realization that most of the energy available in the magnetic flux beneath the surface already has been liberated in expansion by the time the persistent magnetic field configurations above the surface have formed (Book, 1981). These dynamic flux-emergence problems are very difficult to solve and there has been little incentive to do so since magnetic transients in the low corona have rarely been seen. Though hypothesized, for example, as flare triggers, or as the output of magnetic reconnection, fine-scale magnetic transients have defied attempts to measure them quantitatively, again because of the small length and time scales involved. And while there exist many examples of the emergence of a bipolar flux loop through the solar surface, observation of the submergence or disappearance of such loops is generally ambiguous. Only recently a few puzzling cases of "unipolar" flux emergence or decay have been seen in which the other polarity footpoint is missing or hidden.

In the past the dynamic nonlinear phenomena arising during flux emergence have taken a back seat to studies of phenomena which can be represented as small perturbations about a stationary plasma-magnetic field equilibrium. Magnetic loops and arches, for example, persist on the Sun for periods much longer than an Alfvén or sonic transit time along the magnetic flux loop and hence suggest linear stability analyses of a steady state. It is natural to assume that what one sees most of the time is more important than transient events which are inferred but not observed. Further, analysis of linearized problems is much easier than problems where nonlinearities must be treated explicitly. For these reasons linear mechanisms such as unstable MHD equilibria, waves, magnetic reconnection via resistive tearing, and current-driven plasma instabilities have all received close consideration as possible explanations for flares, coronal heating, the solar wind, and other dynamic solar phenomena.

The second observational issue, detecting weak, large-scale patterns, affects our understanding both of flux disappearance and of the magnetic dynamo-solar cycle. The relationship between the solar cycle and large-scale magnetic flux dispersal across the solar disk resulting from the combined action of the Sun's differential rotation and random walk diffusion has been recognized for two decades (Leighton, 1964). More recently, poleward meridional flows have been invoked to help explain this large-scale motion of flux and the timing of the polar field reversal. Measurements of meridional flow are difficult, however, and require extraction of  $\sim 10$  m/s flows over a large area where km/s transient flows are present almost everywhere. Correlating these flows with magnetic fields near the Sun's poles is further complicated by limb effects which degrade observations and by magnetic fields which lie almost perpendicular to the line of sight.

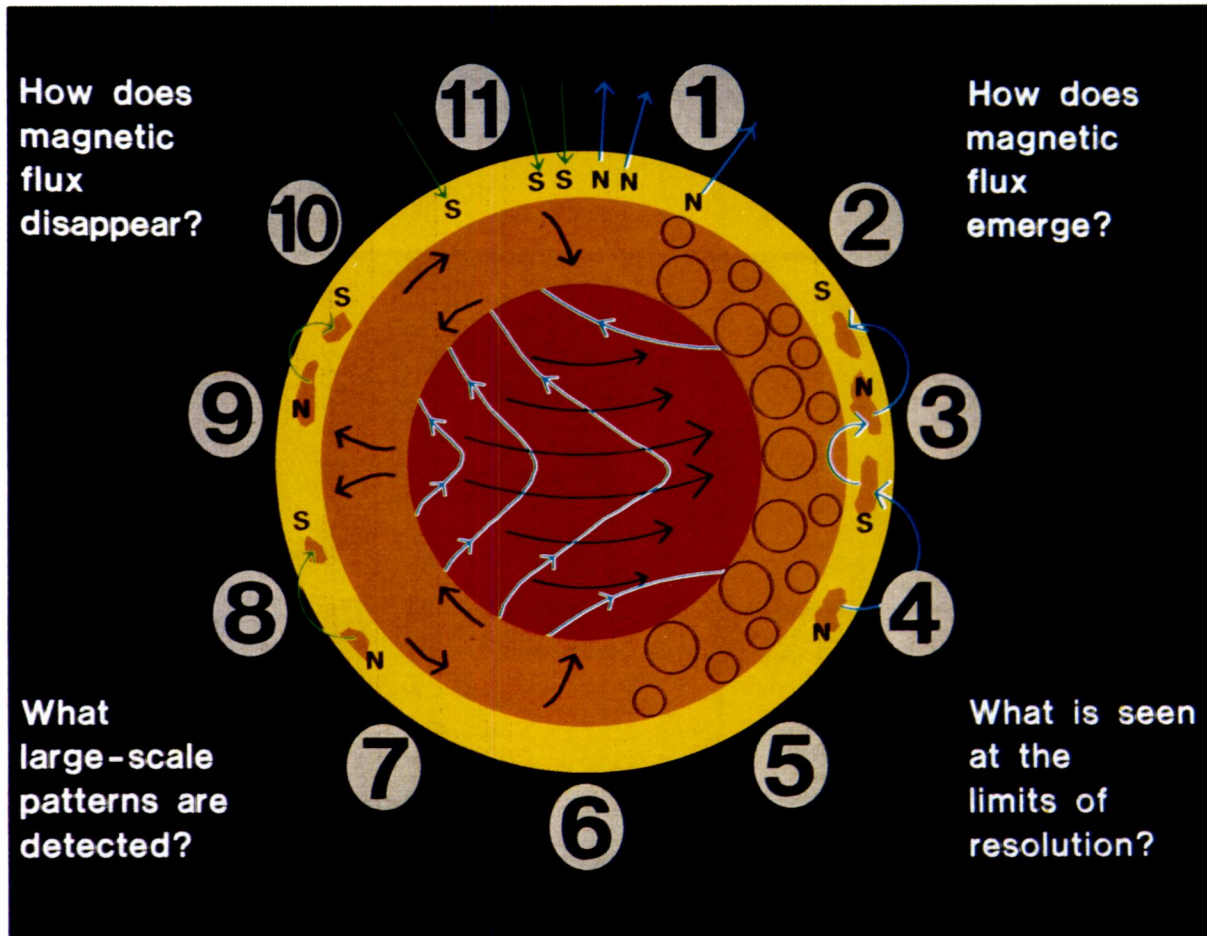
With roughly two complete solar cycles of regular full-disk magnetic data now available, solar physicists are beginning to understand the evolution of the large-scale, long-lived magnetic patterns. These features must now be related to the large convection cells beneath the solar surface, and hence to the solar dynamo, which may be driven in part by this "giant granulation." Again, space-based observation will be the key to achieving the continual coverage in both area and duration that is needed to average out short period oscillations and transients effectively.

Even if accurate, reliable simulation models are eventually developed to predict important magnetic events on the Sun such as flares and the formation of high-speed solar wind streams, the computational forecasts, like terrestrial weather forecasts, will only be as good as the initial conditions inserted into the models from observations. The deviations of the predictions of a weather model from reality over a short period can often be traced to a data net which is too coarse or a gap between subsequent observations which is too large. Because the magnetic fields in the interplanetary medium spiral away from the Sun, the Earth and its magnetosphere can be greatly influenced by events occurring on the far side of the Sun. Since any given point on the solar surface is turned away from the Earth for a couple of weeks at a time, it is clear that some capability to see the other side of the Sun will eventually be required to allow adequate prediction of solar flares, prominences, coronal transients, and high-speed solar wind streams.

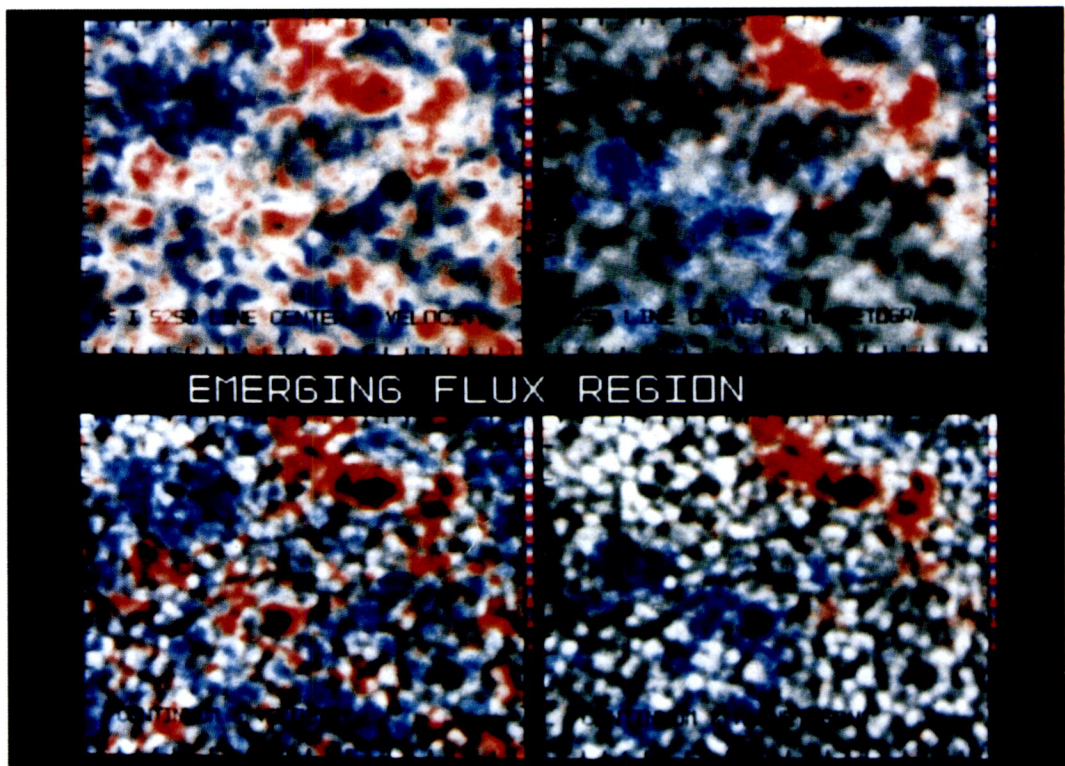
We conclude with a discussion of Figures 3-6 and 3-7, which review, in the context of the solar cycle, the location and timing of the various phenomena discussed in this chapter. Figure 3-6 is a schematic cutaway of the Sun, showing the nonconvective core in red, the convection zone in orange, and a small surface strip in yellow representing the visible photosphere. A possible meridional circulation pattern in and the differential rotation of the convection zone are indicated by the black arrows, and the black circles represent the convective cell structures in the atmosphere. The eleven years of the sunspot cycle are arranged around the periphery. Magnetic field lines are sketched for times early (blue) and late (green) in the cycle. At Year 1, the dipole field is positive at the north pole. As the cycle progresses, field lines are stretched out by the differential rotation (blue lines on the red core), and by Year 3 some of the amplified flux begins to rise buoyantly to the surface, forming large active regions. These regions are oriented such that their meridional dipole moments oppose the global dipole moment; as the regions decay and disperse, their residual flux gradually annihilates the initial polar fields. Around activity maximum (Year 6), the global dipole moment vanishes and then reverses. As new flux continues to emerge with the same orientation, though in dwindling amounts, the polar fields build in strength with the addition of the residual active region flux (Year 9). By the end of the cycle, Year 11, the dipole field has roughly been restored in magnitude but reversed in polarity; sunspot activity is at a minimum; and a new cycle is ready to begin.

Figure 3-7 shows a projection of the entire solar surface where both latitude  $\phi$  and longitude  $\theta$  are measured linearly. The north pole is the entire top edge of the figure and the south pole is the bottom edge. The equator is shown as a thin dashed line. The figure should be wrapped around on itself periodically by merging the right and left edges. A single schematic field line is shown entering the Sun at the south pole, wrapping almost three times around the core, crossing the equator at the righthand edge of the figure, unwrapping in the Northern Hemisphere, and exiting at the north pole. Differential rotation has caused the generation of stronger toroidal flux from the stretching of the north-south dipole field. Numerous bipolar regions are shown with the following polarity regions at slightly higher latitudes than the leading regions. Since the following regions are also of opposite polarity to the nearest polar field, the leading flux tends to cancel more across the equator leaving the unmatched following polarity to eventually cancel and then reverse the polar fields. The subsurface ends of these following field regions must reconnect up beneath the surface, perhaps as indicated by the semicircular arrows near the center of the figure. When the two leading spots in the left center of the figure annihilate, the flux which trails the following regions to the left of center is reconnected beneath the surface. Subsequent migration of some of the trailing spot flux to the poles leaves this subsurface seed flux of the opposite polarity which can then be amplified by the solar dynamo, here shown acting through differential rotation. This view points up the important global role being played by reconnection. As a minimum, subsurface reconnection must be large enough to change the direction of the solar dipole every 11 years.





- 3-6. A schematic flux tube life cycle - the solar clock. The 11-year cycle is arranged around the periphery of a cutaway Sun showing the convection zone (orange) in cross section from the photosphere (yellow) down to the core (red). The large-scale circulation (black arrows) and turbulent convection cells (black circles) combine to drive the Sun's magnetic activity, including the reversal of the global dipole field between the beginning of the cycle (blue lines) and the end of the cycle (green lines).



3-7. Geometry of field lines beneath the solar surface, showing the formation of leading and following active regions, complexes of activity, and the role of reconnection beneath the surface near the solar equator. This figure shows a complementary view of the subsurface magnetic flux also depicted in Figure 3-6. Bipolar magnetic regions generally emerge with a leading flux group appearing to the right of the corresponding following group, at a lower latitude, and with the same sign of magnetic flux as at the corresponding pole.

## VI. REFERENCES

- Abarbanell, C., and H. Wohl, Solar rotation velocity as determined from sunspot drawings of J. Hevelaius in the seventeenth century, Solar Phys., **70**, 197, 1981.
- Babcock, H. W., The topology of the Sun's magnetic field and the 22-year solar cycle, Ap. J., **133**, 572, 1961.
- Babcock, H. W., and H. D. Babcock, The Sun's magnetic field, 1952-1954, Ap. J., **121**, 349, 1955.
- Beckers, J. M., Observations of solar rotation and meridional flows at the Sacramento Peak Observatory, in Workshop on Solar Rotation, edited by G. Belvedere and L. Paterno, p.199, Univ. of Catania, 1979.
- Beckers, J. M., and E. H. Schroter, The intensity, velocity and magnetic structure of a sunspot region, I. Observational technique; properties of magnetic knots, Solar Phys., **4**, 142, 1968a.
- Beckers, J. M., and E. H. Schroter, On the relation between the photospheric intensity, velocity and magnetic fields, Solar Phys., **4**, 165, 1968b.
- Beckers, J. M., and W. R. Taylor, Some comments on the limb shift of solar lines, III. Variation of limb shift with solar latitude, across plages, and across supergranules, Solar Phys., **68**, 41, 1980.
- Book, D. L., A mechanism for heating the solar corona, Comments Plasma Phys. Controlled Fusion **6**, 193, (1981).
- Boris, J. P., N. R. Sheeley, Jr., T. R. Young, Jr., C. R. DeVore, and K. L. Harvey, Meridional flows and magnetic flux transport on the Sun, Bull. Am. Astron. Soc., **15**, 701, 1983.
- Bumba, V., Beginning stages of local magnetic field formation, Bull. Astron. Inst. Czech., **32**, 129, 1981.
- Bumba, V., and R. Howard, Large-scale distribution of solar magnetic fields, Ap. J., **141**, 1502, 1965.
- Cattaneo, F., C. A. Jones, and N. O. Weiss, Periodic and aperiodic behavior in stellar dynamos, to appear in Solar and Stellar Magnetic Fields: Origins and Coronal Effects, IAU Symp. No. 102, edited by J. O. Stenflo, D. Reidel, Hingham,, 1983.
- Chapman, G., and N. R. Sheeley, Jr., The photospheric network, Solar Phys., **5**, 442, 1968.
- Davis, J.M., X-ray bright points and the solar cycle: Further results and predictions, Solar Phys., in press, 1983.
- Deslandres, H., Organisation de l'enregistrement quotidien de la Chromosphere entiere du Soleil a l'observatoire Meudon, premiers resultants, Comptes Rendus, **129**, 1222, 1899.
- Duvall, T. L., Jr., Large-scale solar velocity fields, Solar Phys., **63**, 3, 1979.
- Eddy, J. A., Historical and arboreal evidence for a changing Sun, in The New Solar Physics, AAAS Selected Symp. No. 17, edited by J. A. Eddy, p.11, Westview Press, Boulder, 1978.

- Eddy, J. A., P. A. Gilman, and D. E. Trotter, Solar rotation during the Maunder Minimum, Solar Phys., **46**, 3, 1976.
- Gaizauskas, V., K. L. Harvey, J. W. Harvey, and C. Zwaan, Large-scale patterns formed by solar active regions during the ascending phase of cycle 21, Ap. J., **265**, 1056, 1983.
- Galloway, D. J., M. R. E. Proctor, and N. O. Weiss, Magnetic flux ropes and convection, J. Fluid Mech., **87**, 243, 1978.
- Galloway, D. J., and N. O. Weiss, Convection and magnetic fields in stars, Ap. J., **243**, 945, 1981.
- Gilman, P. A., Global circulation and the solar dynamo, in *The Sun as a Star*, edited by S. Jordan, p. 231, NASA SP-450, NASA, Washington, 1981.
- Gilman, P. A., The solar dynamo: Observations and theories of solar convection, global circulation, and magnetic fields, to appear in *Physics of the Sun*, edited by P. Sturrock, University of Chicago Press, Chicago, 1983.
- Gilman, P. A., and R. Howard, Variations in solar rotation with the sunspot cycle, submitted to Ap. J., 1983.
- Glatzmaier, G. A., Numerical simulations of stellar convective dynamos, I. The model and the method, submitted to J. Comp. Phys., 1983.
- Golub, L., X-ray bright points and the solar cycle, Phil. Trans. R. Soc. Lond., **A 297**, 595, 1980.
- Golub, L., J. M. Davis, and A. S. Krieger, Anticorrelation of X-ray bright points with sunspot number 1970 - 1978, Ap. J., **229**, 45, 1979.
- Hill, F., J. Toomre, and L. November, Variability in the power spectrum of solar five-minute oscillations, Solar Phys., **82**, 411, 1983.
- Howard, R., A possible variation of the solar rotation with the activity cycle, Ap. J., **210**, L159, 1976.
- Howard, R., The rotation of the Sun, Rev. Geophys. Space Phys., **16**, 721, 1978.
- Howard, R., Evidence for large-scale velocity features on the Sun, Ap. J., **228**, L45, 1979.
- Howard, R., and B. J. LaBonte, Surface magnetic fields during the solar activity cycle, Solar Phys., **74**, 131, 1981.
- Howard, R., and J. O. Stenflo, On the filamentary nature of solar magnetic fields, Solar Phys., **22**, 402, 1972.
- Knobloch, E., The diffusion of scalar and vector fields by homogeneous stationary turbulence, J. Fluid Mech., **83**, 129, 1977.
- Knobloch, E., and R. Rosner, On the spectrum of turbulent magnetic fields, Ap. J., **247**, 300, 1981.

- Kraichnan, R. H., Diffusion of passive-scalar and magnetic fields by helical turbulence, J. Fluid Mech., **77**, 753, 1976.
- LaBonte, B. J., and R. Howard, Torsional waves on the Sun and the activity cycle, Solar Phys., **75**, 161, 1982a.
- LaBonte, B. J., and R. Howard, Solar rotation measurements at Mount Wilson III. Meridional flow and limbshift, Solar Phys., **80**, 361, 1982b.
- Leighton, R. B., Observations of solar magnetic fields in plage regions, Ap. J., **130**, 366, 1959.
- Leighton, R. B., Transport of magnetic fields on the Sun, Ap. J., **140**, 1547, 1964.
- Leighton, R. B., A magnetokinematic model of the solar cycle, Ap. J., **156**, 1, 1969.
- Leighton, R. B., R. W. Noyes, and G. W. Simon, Velocity fields in the solar atmosphere, I. Preliminary report, Ap. J., **135**, 474, 1962.
- Livingston, W., and T. L. Duvall, Jr., Solar rotation, 1966-1978, Solar Phys., **61**, 219, 1979.
- Livingston, W., and J. Harvey, The Kitt Peak magnetograph, IV. 40-channel probe and the detection of weak photospheric fields, in Solar Magnetic Fields, IAU Symp. No. 43, edited by R. Howard, p.51, D. Reidel, Dordrecht, 1971.
- Low, B. C., Eruptive solar magnetic fields, Ap. J., **251**, 352, 1981.
- Martin, S. F., and K. L. Harvey, Ephemeral active regions during solar minimum, Solar Phys., **64**, 93, 1979.
- Mayfield, E., and G. Chapman, Magnetic flux changes associated with the solar flares of August 1972, Solar Phys., **70**, 351, 1981.
- McIntosh, P.S., Annotated Atlas of H $\alpha$ Synoptic Charts, UAG Report No. 70, NOAA World Data Center A for Solar-Terrestrial Physics, Boulder, 1979.
- Meyer, F., H. U. Schmidt, G. W. Simon, and N. O. Weiss, Bouyant magnetic flux tubes in supergranules, Astron. Astrophys., **76**, 35, 1979.
- Mosher, J. M., The Magnetic History of Solar Active Regions, Ph.D. thesis, California Institute of Technology, 1977.
- November, L.J., J. Toomre, K.B. Gebbie, and G.W. Simon, The detection of mesogranulation on the Sun, Ap. J., **245**, L123, 1981.
- Parker, E. N., The formation of sunspots from the solar toroidal field, Ap. J., **121**, 491, 1955.
- Parker, E. N., Kinematical hydromagnetic theory and its application to the low solar photosphere, Ap. J., **138**, 552, 1963.
- Parker, E. N., Hydraulic concentration of magnetic fields in the solar photosphere, I. Turbulent pumping, Ap. J., **189**, 563, 1974a.

- Parker, E. N., Hydraulic concentration of magnetic fields in the solar photosphere, II. Bernoulli effect, Ap. J., **190**, 429, 1974b.
- Parker, E. N., The generation of magnetic fields in astrophysical bodies, X. Magnetic buoyancy and the solar dynamo, Ap. J., **198**, 205, 1975.
- Parker, E. N., The generation of magnetic fields in astrophysical bodies, XI. The effect of magnetic buoyancy on the growth and migration of dynamo waves in the Sun, Ap. J., **215**, 370, 1977.
- Parker, E. N., Hydraulic concentration of magnetic fields in the solar photosphere, VI. Adiabatic cooling and concentration in downdrafts, Ap. J., **221**, 368, 1978.
- Parker, E. N., *Cosmical Magnetic Fields, Their Origin and Their Activity*, Oxford: Clarendon Press, New York, 1979a.
- Parker, E. N., Sunspots and the physics of magnetic flux tubes, I. The general nature of the sunspot, Ap. J., **230**, 905, 1979b.
- Patterson, A., and H. Zirin, Transient magnetic field changes in flares, Ap. J., **243**, L99, 1981.
- Perez Garde, M., M. Vazquez, H. Schwan, and H. Wohl, Large-scale solar motions as determined by Doppler shift measurements using a linear photodiode array, Astron. Astrophys., **93**, 67, 1981.
- Proctor, M.R.E., and N.O. Weiss, Magnetoconvection, Rep. Prog. Physics, **45**, 1317, 1982.
- Ramsey, H. E., S. A. Schoolman, and A. M. Title, On the size, structure, and strength of the small-scale solar magnetic field, Ap. J., **215**, L41, 1977.
- Rosner, R., L. Golub, B. Coppi, and G. S. Vaiana, Heating of the corona by anomalous current dissipation, Ap. J., **222**, 317, 1978.
- Schmitt, J. H. M. M., and R. Rosner, Doubly-diffusive magnetic buoyancy instability in the solar interior, Ap. J., **265**, 901, 1983.
- Schussler, M., On buoyant magnetic flux tubes in the solar convection zone, Astron. Astrophys., **56**, 439, 1977.
- Schussler, M., Magnetic buoyancy revisited: Analytical and numerical results for rising flux tubes, Astron. Astrophys., **71**, 79, 1979.
- Sheeley, N. R., Jr., Measurements of solar magnetic fields, Ap. J., **144**, 723, 1966.
- Sheeley, N. R., Jr., Observations of small-scale solar magnetic fields, Solar Phys., **1**, 171, 1967.
- Sheeley, N. R., Jr., private communication, 1982a.
- Sheeley, N. R., Jr., The coronal field lines of an evolving bipolar magnetic region, Ap. J., **255**, 316, 1982b.
- Sheeley, N. R., Jr., J. D. Bohlin, G. E. Brueckner, J. D. Purcell, V. Scherrer, and R. Tousey, XUV Observations of coronal magnetic fields, Solar Phys., **40**, 103, 1975.

- Sheeley, N. R., Jr., J. P. Boris, T. R. Young, Jr., C. R. DeVore, and K. L. Harvey, A quantitative study of magnetic flux transport on the Sun, to appear in *Solar and Stellar Magnetic Fields: Origins and Coronal Effects*, IAU Symp. No. 102, edited by J. O. Stenflo, D. Reidel, Hingham, 1983.
- Simon, G. W., and R. B. Leighton, Velocity fields in the solar atmosphere, III. Large-scale motions, the chromospheric network, and magnetic fields, Ap. J., **140**, 1120, 1964.
- Simon, G. W., and R. W. Noyes, Observations of the coronal network, in *Solar Magnetic Fields*, IAU Symp. No. 43, edited by R. Howard, p.663, D. Reidel, Dordrecht, 1971.
- Simon, G. W., and N. O. Weiss, Supergranules and the hydrogen convection zone, Z. Astrophys., **69**, 435, 1968.
- Spruit, H. C., Magnetic flux tubes, in *The Sun as a Star*, edited by S. Jordan, p. 385, NASA SP-450, NASA, Washington, 1981.
- Stenflo, J. O., Small-scale solar magnetic fields, in *Basic Mechanisms of Solar Activity*, IAU Symp. No. 71, edited by V. Bumba and J. Kleczek, p.69, D. Reidel, Hingham, 1976.
- Stix, M., Dynamo theory and the solar cycle, in *Basic Mechanisms of Solar Activity*, IAU Symp. No. 71, edited by V. Bumba and J. Kleczek, p.367, D. Reidel, Hingham, 1976.
- Stix, M., Theory of the solar cycle, Solar Phys., **74**, 79, 1981.
- Svalgaard, L., T. L. Duvall, Jr., and P. H. Scherrer, The strength of the Sun's polar fields, Solar Phys., **58**, 225, 1978.
- Topka, K., and T. Tarbell, High spatial resolution magnetic observations of an active region, to appear in *Symposium on Small-Scale Dynamical Processes in Solar and Stellar Atmospheres*, edited by S. L. Keil, 1983.
- Topka, K., R. Moore, B. J. LaBonte, and R. Howard, Evidence for a poleward meridional flow on the Sun, Solar Phys., **79**, 231, 1982.
- Tsinganos, K. C., Sunspots and the physics of magnetic flux tubes, X. On the hydrodynamic instability of bouyant fields, Ap. J., **239**, 746, 1980.
- Vaughan, A. H., The Mount Wilson program for stellar activity cycles, to appear in *Solar and Stellar Magnetic Fields: Origins and Coronal Effects*, IAU Symp. No. 102, edited by J. O. Stenflo, D. Reidel, Hingham, 1983.
- Wallenhorst, S. G., and R. Howard, On the dissolution of sunspot groups, Solar Phys., **76**, 203, 1982.
- Wallenhorst, S. G., and K. P. Topka, On the disappearance of a small sunspot group, Solar Phys., **81**, 33, 1982.
- Wilson, P., On the MHD approximation for photospheric magnetic fields, Ap. J., submitted 1983.
- Wilson, P., and G. Simon, Flux changes in small magnetic regions, Ap. J., in press 1983.



- Wu, S., Y. Hu, K. Krall, M. Hagyard, and J. Smit, Jr., Modeling of energy buildup for a flare-productive region, Solar Phys., in press, 1983.
- Yoshimura, H., Nonlinear astrophysical dynamos: The solar cycle as a nonlinear oscillation of the general magnetic field driven by the nonlinear dynamo and the associated modulation of the differential rotation-global convection system, Ap. J., **220**, 692, 1978a.
- Yoshimura, H., Nonlinear astrophysical dynamos: Multiple-period dynamo wave oscillations and long-term modulations of the 22 year solar cycle, Ap. J., **226**, 706, 1978b.
- Yoshimura, H., The solar cycle period-amplitude relation as evidence of hysteresis of the solar-cycle nonlinear magnetic oscillation and the long-term (55 year) cyclic modulation, Ap. J., **227**, 1047, 1979.
- Zirin, H., The magnetic structure of plages, in Chromospheric Fine Structure, IAU Symp. No. 56, edited by R. G. Athay, p.161, D. Reidel, Dordrecht, 1974.
- Zwaan, C., Solar magnetic structure and the solar activity cycle: Review of observational data, in The Sun as a Star, edited by S. Jordan, p. 163, NASA SP-450, NASA, Washington, 1981.



## CHAPTER 4

# MHD WAVES AND TURBULENCE IN THE SUN AND INTERPLANETARY MEDIUM

### WORKING GROUP MEMBERS

Aaron Barnes, Chairman  
*NASA/Ames Research Center*

Melvyn Goldstein  
*NASA/Goddard Space Flight Center*

Joseph Hollweg  
*University of New Hampshire*

John Mariska  
*Naval Research Laboratory*

William Matthaeus  
*College of William and Mary*

Charles Smith  
*University of New Hampshire*

Edward Smith  
*Jet Propulsion Laboratory*

Robert Stein  
*Michigan State University*

George Withbroe  
*Center for Astrophysics*

Richard Woo  
*Jet Propulsion Laboratory*

**CHAPTER 4**  
**MHD WAVES AND TURBULENCE IN THE SUN AND INTERPLANETARY MEDIUM**

I.	Introduction	4-3
II.	Global Oscillations of the Sun	4-4
III.	Observations Related to Waves or Turbulence in the Solar Atmosphere	4-5
IV.	Local Waves in the Solar Atmosphere: Theoretical Considerations	4-9
V.	Interplanetary Hydromagnetic Fluctuations	4-12
VI.	Recent Studies of the Interplanetary Plasma Based on Turbulence Theory	4-13
VII.	Effects of Waves and Turbulence of the Solar Wind	4-17
VIII.	References	4-18

## I. INTRODUCTION

The solar corona is the hot, tenuous outer atmosphere of the Sun. The morphology of the corona is complex, but in the simplest view may be considered a composite of two kinds of regions: the regions of closed magnetic field lines where relatively dense plasma is confined, and regions of open field lines in which the coronal plasma expands freely into interplanetary space. The expanding plasma is accelerated to supersonic speeds within a few solar radii of the Sun's surface, and continues outward through the solar system as the solar wind. Direct observations of the solar wind by the spacecraft Pioneer 10 establish that the flow remains supersonic at least as far as the orbit of Neptune ( $\sim 30$  astronomical units). Hence, the outer plasma environment of all the planets is governed by the solar wind. Theoretical considerations of the interaction of the wind with the local interstellar environment suggest that the plasma will shock down to subsonic flow at some distance in the range  $\sim 50$ - $100$  AU.

The temperature of the coronal plasma varies from place to place, but is generally well above  $10^6$ K, in contrast with the photospheric temperature of  $\sim 5 \times 10^3$ K. Above the photosphere is a somewhat hotter  $\sim 1000$  km thick layer called the chromosphere, then a thin "transition region" in which the temperature rises and density drops to coronal levels. It is not yet known which processes produce the high coronal temperatures although a number of theories have been offered. Because of its high temperature, the corona is an excellent thermal conductor, and heat conducted along open magnetic field lines provides some energy for acceleration of the solar wind. However, it is now well established that classical collisional thermal conduction is inadequate to account for the energy in the observed solar wind, especially in the high-speed streams common near solar minimum. Some additional mechanism(s) for acceleration and heating over an extended region ( $\sim 10$ - $20$  solar radii or more) seems to be required.

Most theories of acceleration or heating in the chromosphere, transition region, corona and solar wind have invoked dissipation of waves or turbulence generated by the solar convective zone or electric currents associated with the evolution of large-scale magnetic structures. However, it is not yet established which processes are dominant. In either case, dissipation may involve the production of turbulent regions.

In situ observations of the solar wind plasma and magnetic field show substantial fluctuations on all observed length and time scales. Variations on the largest scale are due to rotation of large-scale coronal structure under the point of observation, combined with the dynamical evolution of this structure as the wind flows outward. Structures on smaller scales down to the ion gyroradius ( $\sim 100$  km at 1 AU) are thought to be manifestations of hydromagnetic (also called magnetohydrodynamic or MHD) waves or turbulence. Fluctuations on these scales may be of solar origin, or they may be internally generated in the interplanetary medium. To the extent that they originate at the Sun, they may be a fossil of the solar wind acceleration process. In addition, hydromagnetic fluctuations may play a significant role in the microphysics of the solar wind, in regulating the flow of the helium component of the wind, and in the transport of energetic charged particles. In addition to their roles in solar and interplanetary dynamics, hydromagnetic waves and turbulence are thought to be common and dynamically important phenomena in other astrophysical and space plasmas, as well as in fusion plasmas. The solar wind is at present our best laboratory for the study of large-amplitude hydromagnetic fluctuations, and the study of MHD waves and turbulence has been a major thrust of interplanetary physics.

This report, based on the deliberations of the working group, is our overview of the current state of knowledge concerning the existence, nature and dynamics of MHD waves and turbulence in the solar atmosphere and interplanetary medium. We first consider remote-sensing observations of global oscillations of the Sun, of nonthermal motions in the chromosphere, transition region, and corona, and their possible interpretations in terms of waves or turbulence. We next consider the region of solar wind acceleration; here coronal imaging, resonance-line spectrometry and radio techniques provide the relevant observations. We then consider fluctuations and discontinuities in the interplanetary medium itself, in which the primary data come from direct sampling of the plasma by spacecraft-borne magnetometers and plasma analyzers. We do not consider microscopic plasma instabilities or waves/turbulence. We do not attempt to provide a comprehensive review of any of these topics; rather our goal has been to assess how successful past attempts to close theory and observation have been, and to obtain a sense of where future advances may lie.

## II. GLOBAL OSCILLATIONS OF THE SUN

One of the most dramatic achievements of solar physics in the last decade was the discovery of the modal structure of the solar oscillations and the beginning of their use to diagnose the Sun's interior. Observations of surface motions resulting from wave modes that propagate throughout the Sun provide an opportunity to measure the Sun's internal structure. These methods are similar to those applied in terrestrial seismology and can also be applied to other stars.

The most extensively studied oscillations have periods in the neighborhood of five minutes. These are sound or p- (for pressure) waves trapped in a cavity between the depth where the waves are refracted back towards the surface (when the horizontal phase speed equals the local sound speed) and the surface where the waves are reflected downward (when the wavelength becomes comparable to the density scale height). The oscillations must have periods longer than three minutes, because shorter period waves are not reflected at the photosphere. The normal modes are those waves which make a circuit of the cavity and return in phase with themselves, that is for which an integral number  $n$  of half-vertical wavelengths fit into the cavity,  $k_z L = [2\pi v/s]L = (n+1/2)\pi$ . For long-lived modes, that can propagate horizontally around the Sun and interfere with themselves, an integral number of horizontal wavelengths must fit around the circumference of the Sun, so  $k_H = \ell/r$ . High angular degree  $\ell$  modes are confined to a shallow cavity beneath the solar surface and thus probe the structure of the convection zone. Modes of decreasing degree  $\ell$  penetrate deeper into the solar interior, with the  $\ell=0$  cavity extending all the way to the center (Christensen-Dalsgaard, 1980; Leibacher and Stein, 1981).

Comparison of the observed modal structure with theoretical frequencies calculated from solar models has already shown that the convection zone must be deep, covering approximately the outer third of the Sun, and that the solar helium abundance is  $Y \sim 0.25$ . This helium abundance is consistent with the "big bang" and standard solar models, but inconsistent with low helium models proposed to solve the solar neutrino problem. High degree modes have observed frequencies below those predicted by the standard solar model with mixing length of one pressure scale height. This means that the sound speed in the convection zone must be smaller and the adiabat lower than in the standard model. Hence the super adiabatic gradient must be smaller and convection more efficient, which implies a larger mixing length. This in turn produces a deeper convection zone. The helium abundance can be determined because less helium means more hydrogen, hence more particles per unit mass, and a larger pressure for a

given temperature. Thus the central temperature would be lower to support the weight of the Sun. A lower central temperature means a smaller temperature gradient and lower luminosity, unless the opacity were also reduced, which requires a lower abundance of heavy elements. Such a model is cooler in the center, but hotter in the outer layers than the standard model. This would lead to larger frequencies for high degree modes in contradiction to the observations. Conversely, it would lead to smaller frequencies for low degree modes, because these probe primarily the deep interior, which is cooler in a low helium model (Gough, 1982).

It is possible to extract even more information on the internal structure of the Sun from the observed modal structure of the low degree modes which penetrate into the core by solving an inverse problem. Each mode's frequency is a weighted average of conditions below the visible surface. Since modes of different radial order  $n$  and angular degree penetrate to different depths, the difference between the eigen functions of similar modes will be large only in a thin shell. Such differences can then be used to determine the structure of the Sun in this thin shell. Preliminary application of this method to rotational splitting of modal frequencies in limb observations suggests that the Sun's core is rotating six times faster than its surface. In practice it is found that the linear combinations needed to construct localized differences in modes, when just p-modes are used, sometimes have very large coefficients, which produce errors in the inferred structure an order of magnitude larger than the errors in the observations.

This situation can be greatly improved by also using g-modes, waves driven by the buoyancy force, which can propagate in a convectively stable fluid. If a few of these g- (gravity) modes could be observed, the inversion would be substantially improved, because g-modes have a very different structure than p-modes. The energy in g-modes is concentrated toward the center of the star and decays going through the convection zone, (which makes them hard to detect). This makes the g-modes sensitive to very different aspects of solar structure and greatly increases the reliability of any inversion procedure. Gravity modes have frequencies lower than the maximum of the Brunt-Vaisala frequency in the solar interior, which gives them periods longer than most p-modes. The search for these g-modes, one of which may already have been detected in the 160 minute mode, should be a top priority (Christensen-Dalsgaard, 1980; Leibacher and Stein, 1981).

Both the p- and g-mode spectra are very dense. Why only p-modes with periods near five minutes, and not up to the hour period of the fundamental, have been observed is a vexing theoretical question. Why only a single g-mode is seen is even more puzzling. (Others may have recently been observed.) These are probably issues of mode excitation and damping, which is an area for further investigation.

### III. OBSERVATIONS RELATED TO WAVES OR TURBULENCE IN THE SOLAR ATMOSPHERE

#### CHROMOSPHERE, TRANSITION REGION AND INNER CORONA

While photospheric oscillations are reasonably well understood, further out, in the upper chromosphere, transition region, and inner corona, the observational picture is less comprehensive and our understanding is incomplete.

Optically thin UV lines formed in the upper chromosphere, transition region, and inner corona exhibit broadening in excess of that expected from thermal broadening alone. This additional broadening is due to mass flows associated with small-scale structural features present in these layers, wave motions and turbulence. It is usually characterized by a parameter  $v_t$  which is added to the expression for the Doppler width of the line and is sometimes referred to as the nonthermal or microturbulent velocity. Measurements of average values of this nonthermal velocity in and above the transition region in the quiet Sun show a number of important properties. First, there is little or no change in  $v_t$  from the center to the limb. Second, the value of  $v_t$  increases with increasing temperature of line formation from the upper chromosphere to temperatures near  $10^6\text{K}$ . Typical velocities in the upper chromosphere at a temperature of  $2 \times 10^4\text{K}$  are about 10 km/s, while in the transition region at  $2 \times 10^5\text{K}$   $v_t$  is about 25 km/s (see Figure 4-1). The nonthermal broadening is however always less than or comparable to the local sound speed. Third, the nonthermal broadening varies on time scales from minutes to hours at any location and on spatial scale down to 1 arcsec (725 km). These fluctuations generally become larger at smaller spatial scales.

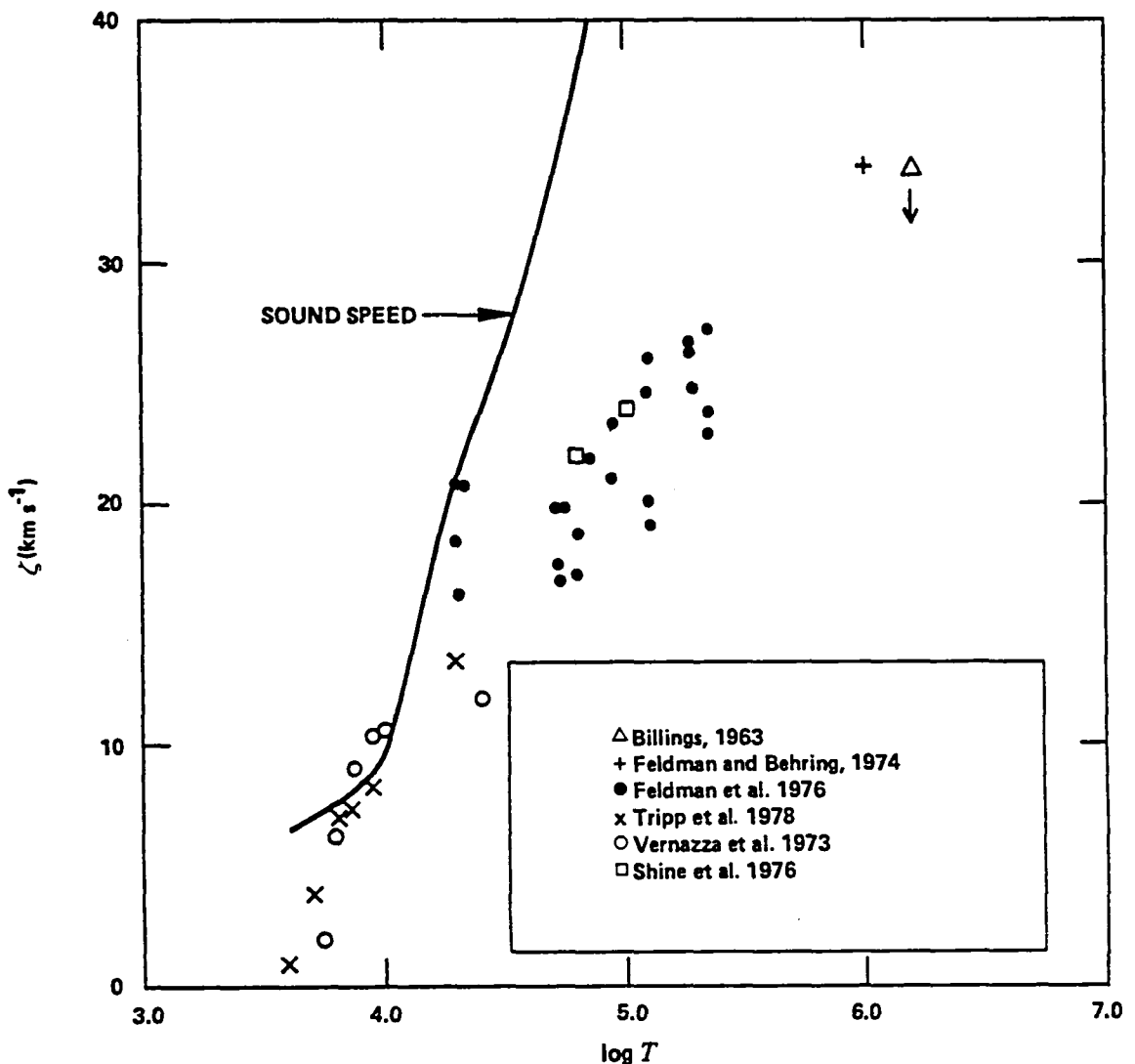


Figure 4-1. Microturbulent velocity as a function of temperature in the solar atmosphere. (From Athay, 1981.)

Analyses of time series of profiles of optically thin emission lines provide additional information on the relation between the observed properties of the line profile and waves and turbulence. Athay and White (1978) have analyzed such time series for SiII emission lines formed in the chromosphere and found evidence for wave motions. In a similar analysis (Athay and White, 1979) for CIV emission lines formed in the transition region at  $10^5\text{K}$ , they found statistically significant wave motions in about 20% of the observations. Bruner (1981) performed a similar analysis and found that the data are consistent with the existence of acoustic or magnetic waves in the transition region, but that downward propagating waves are seen as frequently as upward propagating waves. The magnitude of this nonthermal broadening puts an upper limit on the wave velocity amplitudes, such that only waves propagating at speeds close to the Alfvén speed can transport enough energy through the transition region to heat the corona.

Nonthermal motions are observed well into the inner regions of the corona (e.g., Cheng et al., 1979). They are thus potentially related to the heating and acceleration of the solar wind and possibly some of the wave phenomena observed there. The detailed properties of the nonthermal motions in the inner corona ( $r < 1.1 R_\odot$ ) are, however, known even less well than those of the transition region summarized above. About all that can be stated currently is that the nonthermal velocity is between 10 and 25 km/s and appears to be relatively constant with height out to about  $2 \times 10^4$  km above the limb.

From an observational point of view, our understanding of nonthermal motions in the upper chromosphere, transition region, and low corona is only beginning to emerge. The gross properties at relatively large spatial scales are clear, but studies at spatial scales arcsec and below are just beginning (e.g., Dere et al., 1983). For example bursts of turbulent motion have been seen in UV lines formed near  $10^5\text{K}$  at a rate of  $750\text{ s}^{-1}$  for the entire Sun. Jets of upward moving chromospheric gas called spicules have long been observed in the solar atmosphere, but their nature is still unclear. High resolution UV imaging combined with spectral observations would help clarify what spicules are and do. Such capabilities will be provided by the Solar Optical Telescope. One suggestion is that spicules are produced by a shock wave propagating along a magnetic flux tube lifting the chromospheric material up into the corona. Such a model spicule will oscillate up and down as gravity pulls the plasma down and succeeding shocks from the wake of the initial impulse travel along the spicule and lift it up again, while also slowly heating it. Gas moving up at very high speeds ( $\sim 400\text{ km s}^{-1}$ ) has recently been observed in the UV from ions existing at temperatures near  $10^5\text{K}$ . The role of jets in the energy and mass balance of the corona and solar wind is unclear and needs further investigation. A really complete picture will not emerge until high spatial and temporal resolution emission line profiles can be combined with information about the fine scale structure that makes up much of these regions of the atmosphere. Observations with the UV high resolution telescope and spectrometer on Spacelab 2 and UV spectrometer on the Solar Optical Telescope will provide critical data for the chromosphere and transition region. Instrumentation with similar capabilities (temporal resolution of seconds, spatial resolution  $< \text{arcsec}$ , velocity resolution of few km/sec) at shorter wavelengths are required to probe the upper transition region and inner corona. High resolution observations at radio wavelengths with the VLA can also provide information.

## SOLAR WIND ACCELERATION REGION

The theory of the expansion of the solar corona to form the solar wind, although well-understood in the whole, has facets which are poorly understood in detail. One of these facets is the mechanism by which energy is supplied to power the supersonic expansion. Although it is widely believed that waves originating near the solar surface transport energy to heat the outer corona, detailed testing of competing models of this heating will probably not be possible until in situ measurements can be made in the acceleration region. In the meantime, however, indirect but useful information on the physical state of the acceleration region can be

obtained through remote sensing techniques. In the region above  $1.5 R_{\odot}$  observations of electron and resonantly scattered radiation can provide measurements of electron and ion temperatures, electron and ion densities, solar wind velocities, chemical abundances and upper limits on nonthermal velocities. Radio observations of scattering by electron density irregularities using either natural or spacecraft radio signals can provide measurements of the bulk flow of the solar wind, solar wind velocity fluctuations, electron density and magnetic field fluctuations in the acceleration region close to the Sun.

At levels in the corona beyond  $1.2R$  there are only very limited measurements of spectral line profiles that can provide constraints on the magnitude of nonthermal velocities that could be produced by waves. There is no reliable spectroscopic evidence for the existence of a steady flux of waves, although, as we discuss below, some measurements are consistent with the presence of Alfvén waves in a coronal hole. However, transient phenomena have been observed that can produce outward propagating disturbances, namely coronal transients.

The available observations beyond a few tenths  $R$  above the solar surface consist primarily of eclipse measurements of spectral line profiles of heavy ions, such as the Fe XIV 5303 line and rocket measurements of the resonantly scattered hydrogen Lyman alpha line. These provide reliable upper limits on the magnitude of rms nonthermal velocities near  $r=1.5R$ . For quiet regions of the corona the width of the Fe XIV line corresponds to velocities between 20 and 40 km/sec with a mean value of about 30 km/sec (Liebenberg et al., 1975). This is an upper limit for the magnitude of the nonthermal velocities in the line of sight (which is in the transverse direction, perpendicular to the radial direction because the observations are made at the limb). It is an upper limit because the line width includes a thermal component which is 17 km/sec for Fe XIV at  $10^6\text{K}$  and 24 km/sec at  $2 \times 10^6\text{K}$ . Since temperatures above  $10^6$  are required for Fe XIV to be observed at this height ( $r \approx 1.5 R_{\odot}$ ), this suggests that the mean nonthermal velocity there is approximately 25 km/sec or less. Hence, any transverse waves such as Alfvén waves must have an rms velocity amplitude of less than or equal to this value at  $r \approx 1.5R$  in "quiet" regions of the corona. (By quiet we mean outside of coronal holes, since the Fe XIV line is likely to be too faint to be observed with present instrumentation at these heights in coronal holes.)

Beyond  $\sim 1.5 R_{\odot}$  the only measurements of spectral line profiles are those of hydrogen Lyman alpha acquired on several rocket flights (Kohl et al., 1980, 1983; Withbroe et al., 1982b). These data give information about temperatures and velocities between  $1.5$  and  $4 R_{\odot}$ . Measurements in a polar coronal hole place an upper limit of 140 km/sec on the magnitude of nonthermal motions over this height range. Comparisons with a two-fluid model (used to estimate the proton/hydrogen thermal temperature) indicate that the nonthermal velocity at  $4 R_{\odot}$  could be as high as 110 km/sec. These data are consistent with the presence of Alfvén waves whose rms velocity amplitude grows as  $\rho^{-1/4}$  with increasing height in the atmosphere such that the velocity amplitude at  $4 R_{\odot}$  is 110 km/sec.

In the only quiet region observed thus far the Lyman alpha data provide an upper limit to the nonthermal velocity at  $4 R_{\odot}$  of 140 km/sec. This limit increases with decreasing height reaching 200 km/sec at  $1.5 R_{\odot}$ . Since hydrogen has low mass, the Lyman alpha line is easily broadened by thermal motions (a temperature of  $10^6\text{K}$  yields an rms velocity of 130 km/sec), and thus most of the broadening at low heights is likely to be of thermal origin. Thus, at  $1.5 R_{\odot}$  the limit provided by the Fe XIV measurements referred to above is more restrictive.

The mean solar wind speed accelerates from about 150 km/s at  $10 R_{\odot}$  to about 450 km/s at  $30 R_{\odot}$  as shown by multiple-station intensity scintillations with natural (Ekers and Little, 1971; Scott 1978) and spacecraft (Armstrong and Woo, 1981) radio signals. These same measurements show that a turbulent region surrounds the Sun as random velocities representing a significant fraction of the mean are observed within  $20 R_{\odot}$ . Beyond  $20 R_{\odot}$ , there is a rapid decrease in fractional



random velocities. These velocity fluctuations are consistent with the presence of a large flux of MHD waves. Coronagraph images also indicate the presence of random velocity fluctuations; these data await interpretation. Measurements of flow velocities using Lyman alpha intensity Doppler dimming technique provide evidence for subsonic flow out to  $4 R_{\odot}$  in a quiet region of the corona (Withbroe et al., 1982a) and polar region (Withbroe et al., 1982b), a result consistent with the radio measurements. Preliminary results from a polar coronal hole give evidence for supersonic flow at  $\sim 2 R_{\odot}$  (Kohl et al., 1983). The Lyman alpha data also provide measurements of proton temperatures in the same regions. The proton temperatures are either constant ( $T_p = 10^6 K$ ) for the two polar regions and decreasing with increasing radius (quiet region) from  $2.5 \times 10^6 K$  at  $r = 1.5 R_{\odot}$  to  $10^6 K$  at  $r = 4 R_{\odot}$  for the quiet region.

Extensive radio phase scintillation and spectral broadening observations have been used to study the electron density power spectrum over the approximate elongation range  $2 R_{\odot}$  to 1 AU (Woo and Armstrong, 1979). These measurements show that for heliocentric distances beyond  $20 R_{\odot}$ , the one-dimensional electron density spectrum (as would be seen by in situ spacecraft) is well-modeled by a single power law in the frequency range  $10^{-4}$ - $5 \times 10^{-2}$  Hz with a mean spectral index of 1.65. This value is very close to the Kolmogorov value of  $5/3$  for the inertial subrange; this fact at present has no theoretical explanation. Within  $20 R_{\odot}$ , the density spectrum can still be modeled by a single power law over the frequency range  $10^{-3}$ - $10^1$  Hz, but the spectral index decreases to about 1.1. This flattening of the density spectrum has also been observed by Harmon and Coles (1983) in the spectral broadening of planetary radar signals.

The flattening of the density spectrum between 2 and  $20 R_{\odot}$  is presumably associated with energy deposition in the near-Sun region and acceleration of the solar wind. This is the same turbulent region surrounding the Sun where large fractional velocity variations have been found using the multiple-station intensity scintillations.

When the spacecraft signal is linearly polarized, as is the case with Helios, information on magnetic field fluctuations can be deduced from the observed Faraday rotation fluctuations. The interpretation of the Faraday rotation fluctuations, however, is not as straightforward as the aforementioned radio scattering observations. Hollweg et al. (1982) have analyzed the Helios measurements and have shown that the observed Faraday rotation fluctuations cannot be solely due to electron density fluctuations in the corona unless the coronal magnetic field is some five times stronger than that suggested by current estimates. The Faraday rotation fluctuations are consistent with the hypothesis that the Sun radiates Alfvén waves with sufficient energies to heat and accelerate high-speed solar wind streams.

An improved picture of the acceleration region will be established with the aid of (1) multiple-station intensity scintillation measurements of the solar wind using the VLA, (2) additional Faraday rotation fluctuation observations, and (3) observations with the UV coronal spectrometer on Spartan II. More sophisticated remote sensing experiments and in situ measurements (within 0.1 AU) are required to provide additional critical information on the physical conditions in the solar wind acceleration region and to determine the role of waves in plasma heating, energy transport and solar wind acceleration.

#### IV. LOCAL WAVES IN THE SOLAR ATMOSPHERE: THEORETICAL CONSIDERATIONS

The solar atmosphere is a plasma in a gravitational field penetrated by magnetic fields. As a result three restoring forces -- pressure, buoyancy and Lorentz -- act to propagate wave motions. In horizontally homogeneous regions (where the scale length for variations in

atmospheric properties is greater than the wavelength) the waves are internal or body magneto-acoustic-gravity waves. One mode is an Alfvén wave. In a weak magnetic field the other two resemble a magnetic slow mode and an acoustic (high frequency) or internal gravity (low frequency) fast mode. In a strong field these become a magnetic fast mode and an acoustic slow mode propagating along the magnetic field lines. The gravity mode is suppressed (Stein and Leibacher, 1974; Thomas, 1983). A complete description of these waves is still a matter of current study.

The solar atmosphere and its magnetic field are, however, highly structured, on scales ranging from a solar radius down to the current resolution limits (several hundred kilometers). This structure significantly modifies the propagation and dissipation of waves. Where the magnetic field is clumped into flux tubes the waves are tube modes confined to the region of the magnetic flux tube. The spectrum of tube modes is more complex than for body modes, and for the most part unstudied. In the limit of thin vertical flux tubes there are again just three modes -- a torsional Alfvén wave, a longitudinal acoustic (sausage) slow wave, and a transverse magnetic (kink) fast wave. The latter two both have low frequency cutoffs due to the stratification (Spruit, 1982). For a fatter flux tube there also exists a fast magnetic transverse "sausage" wave which propagates only for wavelengths shorter than the tube diameter at speed greater than the tube Alfvén speed but less than the Alfvén speed in the external medium. Roberts et al. (1983) suggest that oscillations associated with type IV radio bursts may be the signature of these waves. Because fast tube waves are dispersive, they can even evolve into solitons (Roberts and Mangeney, 1982).

Turbulent motions in the convection zone will generate most of these wave modes. To generate torsional Alfvén waves requires vertical vorticity in the turbulence on a scale comparable to the size of the flux tube; this process will probably be inefficient. Torsional Alfvén waves can also result from the unwinding of subsurface twists on the flux tubes (Parker, 1982). The longitudinal tube mode is excited by pressure fluctuations and vertical velocities with frequencies above approximately the acoustic cutoff, while the transverse tube mode is excited by horizontal motions (which predominate near the top of the convection zone) above a somewhat lower cutoff frequency. Thermal instabilities will also generate these wave modes. Since there is so much energy in the turbulent motions, a substantial flux of waves is probably generated, even where the efficiency is low. The dependence of energy fed into different wave modes on the surface properties of the Sun and other stars is still only known approximately (Stein, 1981). In addition, impulsive events, such as flares, will generate all these wave modes.

Once the waves are produced, the next question is where do they propagate their energy? In particular, can waves transmit enough energy to the chromosphere and corona to heat these regions? All internal or body modes suffer substantial refraction and reflection at the transition region, where both the sound and Alfvén speed have a large jump. Fast mode waves from the convection zone, in fact, tend to undergo total internal reflection in the chromosphere, and so never reach the corona. Tube and Alfvén waves, on the other hand, always propagate their energy along the magnetic flux tube. They are partially but never totally reflected in the chromosphere and transition region. Whereas in the photosphere structure is primarily due to clumping of the magnetic field, in the corona the magnitude of the field varies significantly only on large spatial scales. Structure in the corona is due both to large density variations and to changes in magnetic field direction, both of which can occur over short spatial scales and can support surface waves. Closed magnetic field loops in the corona can provide a resonant cavity which acts like an interference filter and allows energy to enter the corona in certain frequency bands without reflection (Hollweg, 1983; Zugzda and Locans, 1982). It may also be possible to treat the propagation of these waves in terms of an equivalent LRC circuit (Ionson, 1982), but

there are differences between the wave and circuit analyses (Hollweg, 1983). In addition, dense coronal structures can trap fast "body" waves, in analogy with optical fibers, (Habbal et al., 1979). An important subject of current research is to assess whether Alfvén and fast tube waves with amplitudes consistent with the observed nonthermal velocities can carry interesting energy fluxes into the upper chromosphere and corona.

Once the waves are produced and the energy is propagated into the chromosphere and corona, the next question is how they dissipate their energy? Compressive waves are damped by radiation in the photosphere, so only about 10% of their energy reaches the chromosphere. Ion-neutral friction, electrical resistivity, and viscosity can also dissipate waves in the chromosphere, if their wavelength is short enough. In the corona, heat conduction and viscosity can dissipate short wavelength waves. Because the density decreases with height, the amplitude of all waves tends to grow with height in the atmosphere, in order to conserve energy. This leads to nonlinear dissipation of the waves. Compressive waves dissipate primarily by shock formation. Gravity waves break and produce small scale turbulence. Incompressible magnetic waves dissipate by coupling to other modes which can dissipate. Although internal Alfvén waves propagate with little dissipation, Alfvén and transverse tube waves dissipate efficiently by phase mixing due to the structured medium. This process can be thought of as follows: each magnetic field line oscillates approximately independently and because of the different field strength and density at different locations within the flux tube the Alfvén speed will be different on different field lines. Hence waves propagating along different field lines travel at different speeds and get out of phase, i.e., may get refracted. This builds up small scale transverse structure which dissipates eventually by viscosity even when viscosity is small. (Heyvaerts and Priest, 1983). This process has been observed in fusion machines. It is more efficient on closed flux tubes, where the waves bounce back and forth many times. Shear viscosity transverse to the magnetic field is so small that the dissipation length for this process is larger than the length of loops. However, the increasing structure may drive instabilities, e.g., Kelvin-Helmholtz instability due to large velocity shears, which may trigger the development of turbulence (Hollweg, 1983). There is an important difference between Alfvén and fast magnetic tube waves: Alfvén waves always propagate their energy along the magnetic field. Only the energy propagating along field lines in the region of large perpendicular Alfvén speed gradient dissipate their energy and heat the medium. Fast waves propagate energy across the field, and so concentrate the energy from a wide area into the resonant dissipation region (Ionson, 1978). The quiet corona, composed of large scale, magnetically closed, loops, requires an energy flux of a few  $\times 10^5$  ergs  $\text{cm}^{-2} \text{s}^{-1}$  to heat it.

Coronal holes, from where originates the high speed solar wind streams, require slightly more energy, about  $10^6$  ergs  $\text{cm}^{-2} \text{s}^{-1}$ . The corona above active regions, which consist of short dense loops (presumably confined by the stronger, 100 gauss, magnetic field there) requires the most heating, about  $10^7$  ergs  $\text{cm}^{-2} \text{s}^{-1}$ . It is likely that the heating processes for closed and open field lines are similar, because the energy input required for closed loops is comparable (sometimes greater, sometimes less) to that required for coronal holes.

The study of magnetoacoustic-gravity wave transport and dissipation of energy has barely begun, and their role in heating the chromosphere and corona needs to be elucidated. Aside from waves, other mechanisms that may heat the corona are newly emerging magnetic flux and steady motion of pre-existing magnetic fields which will produce currents that can dissipate. Current observations are unable to distinguish between the various coronal heating mechanisms which have been suggested. If waves transport the energy to heat the corona, we should be able to observe them. Line profiles produced by such motions should be calculated in order to compare with future high-resolution observations from space. High-resolution observations can also help

to distinguish between different dissipation mechanisms. For instance, phase mixing will likely lead to small-scale turbulent motions, while reconnection may produce localized high-speed jets. Unfortunately, although the energy transport and dissipation processes are magnetic, we cannot directly observe the coronal magnetic field.

## V. INTERPLANETARY HYDROMAGNETIC FLUCTUATIONS

The solar wind typically exhibits fluctuations on all observed scales, encompassing the entire hydromagnetic regime (and beyond) (e.g., see the review articles of Barnes (1979) or Hollweg (1978)). Typical time series of (say) a component of the interplanetary magnetic field have a stochastic appearance, and are often described as turbulent. On the other hand, a propagating wave from a stochastic source will also show a random appearing time series. One of the major issues to be settled is the operational definitions of waves and turbulence, and the observations and/or modes of data analysis required to make the distinction.

One of the traditional views of interplanetary irregularities emphasizes the occurrence of discontinuous or abrupt changes in magnetic field direction (often with  $\Delta B \sim B$ ). The interval between successive discontinuities follows a Poisson distribution. The mean time between discontinuities is  $\sim 1$  hour. According to hydromagnetic theory, two fundamentally different types of discontinuities are possible. Tangential discontinuities are static structures along which the sums of the magnetic and plasma pressures on the two scales balance. Rotational discontinuities are basically large amplitude Alfvén waves, propagating structures which transport momentum and energy. Discontinuities provide a test of hydromagnetic theory. In the limit of thin MHD structures only two types should occur and should satisfy a set of specific constraints. Thus far, theoretical expectations have been confirmed although there is a large class of discontinuities that do not appear clearly to be either tangential or rotational. This class may represent observational difficulties rather than a third type of discontinuity.

Power spectra of the magnetic field and plasma parameters usually show a power law over several decades in wavenumber. This behavior is suggestive of the inertial range of turbulence theory. On the other hand, in much of the solar wind (especially the middle and trailing edges of high speed streams) the observed fluctuations behave very much like Alfvén waves (Belcher and Davis, 1971). For ideal Alfvén waves, the magnetic field varies in direction but not magnitude, and the density remains constant. These characteristics are not restricted to plane waves, but may occur in the presence of a tangled, apparently turbulent field. The observed "Alfvénic fluctuations" are reasonably consistent with the description; however, some workers suggest that the small deviations from ideal Alfvénic fluctuations strength are an important clue to the origin and evolution of the wave/turbulence field.

The correlations between variations in velocity and magnetic field associated with the Alfvénic fluctuations indicate that they are propagating outward from the Sun, strongly suggesting a solar or near-solar origin. The fluctuations usually have a fairly well-defined direction of minimum variance, which would be consistent with their interpretation as plane waves. The radial variation of their amplitude with heliocentric distance is not well determined, but the studies that have been made give results consistent with the radial propagation of Alfvén waves, possibly somewhat attenuated by nonlinear processes. The predominance of the Alfvén mode is consistent with the strong Landau damping of magnetoacoustic waves in a hot, collisionless plasma. However, the observed direction of minimum variance is strongly aligned, on the average, with the mean magnetic field direction; this alignment is inconsistent with theoretical predictions based on simple wave propagation.

Hollweg (1982) has pointed out that surface waves may propagate along the copious tangential discontinuities (current sheets) that compartmentalize the solar wind. These Alfvénic surface waves have some but not all properties of ordinary Alfvénic fluctuations, and could conceivably be a significant or even dominant contributor to the observed wave/turbulence field. The theory has not yet been developed to the point that a clear observational test of this possibility has been defined. But it does offer a possible explanation for the observed minimum-variance alignment. In general, the interrelation between current sheets and wave/turbulence has not yet been adequately investigated.

This alignment of minimum variance direction can be understood if the fluctuations are stochastic and three-dimensional, rather than plane waves (Barnes, 1981; Bavassano and Mariani, 1983). This fact provides an argument favoring the turbulence viewpoint. On the other hand, for ideal stochastic Alfvénic fluctuations the Reynolds and Maxwell stresses cancel, so that the nonlinear cascade of energy associated with turbulence should not occur. This consideration, together with the outward propagation of the fluctuations, suggests that the fluctuations are the remnant of a more complicated turbulent cascade originating near the Sun.

Recently there has been a significant effort to describe the interplanetary observations in the language appropriate to, or derived from hydrodynamic turbulence. Because this work is rather new and still unfamiliar to many workers in the field, a special section of this report will be devoted to it. But for the moment, let us view the interplanetary fluctuations as waves (taking account of the simultaneous presence of convected, static structures such as tangential discontinuities). The simplest mode of theoretical description is fluid MHD theory. Unfortunately, this theory is not completely valid in a collisionless plasma, and important effects (such as the dissipation of compressive waves by Landau damping) are omitted. Another complication is the large-amplitude nonlinear character of the interplanetary fluctuations. Our present mathematical tools for dealing with this kind of problem are limited, and progress may require extensive numerical simulation.

## VI. RECENT STUDIES OF THE INTERPLANETARY PLASMA BASED ON TURBULENCE THEORY

Describing a fluid system as turbulent is equivalent to saying that the dynamical fluid variables exhibit complex and essentially non-reproducible behavior as a function of time. This is generally due to nonlinearities in the fluid equations which strongly couple a large number of degrees of freedom. Turbulent systems are usually far from equilibrium and detailed analytically tractable theories do not exist. The solar wind plasma flow and the interplanetary magnetic field carried along with it resemble such a system. In the zero momentum frame, the magnetic and velocity field fluctuations are energetically comparable to the mean magnetic field over length scales of order 1 AU and display much of the complicated behavior expected of turbulence.

The prospect that plasma turbulence techniques may be applied and developed in the context of solar wind studies is attractive. Although the interplanetary medium cannot be controlled as laboratory plasmas might be, it does persist in time in what appears to be a statistically steady state. Furthermore, we have a vast data base of magnetic field and plasma properties along spacecraft trajectories. The availability of these data provides the opportunity to evaluate existing theories and may provide guidelines for the development of new approaches.

Theoretical studies of interplanetary turbulence must utilize appropriately defined ensemble averaged properties of the fields. The connection between statistical theories of turbulence and experimental reality can be made only when it is possible to show that the theoretician's abstract notion of an ensemble has a plausible correspondence to some experimentally implemented averaging procedure. The connection usually sought is that time averages of the products of the turbulent fields at fixed points in space are equivalent, in practice, to ensemble averaging. When this is so, the fluctuations are "ergodic". If time averages are also insensitive to the origin in time, the fluctuations are statistically stationary.

Consider a component of the magnetic field  $B(t)$  at a fixed point in space. Frequency spectra and two-time correlation functions indicate that  $B(t)$  consists of a continuum of frequency components superimposed on coherent signals with periods of the solar rotation period and its first few harmonics. The signals commensurate with the solar rotation reflect sector structure, stream structure and other nearly periodic phenomena.

Stationarity of the average field  $\langle B(t) \rangle = a$ , and second order moments such as the correlation function  $R(\tau) = \langle B(t)B(t+\tau) \rangle$  is the property that the ensemble average denoted by  $\langle \dots \rangle$  does not depend on  $t$ . To determine the extent to which this property is reflected in the data, Matthaeus and Goldstein (1982a) considered sequences of time averages of the field covariances of the same. According to the ergodic theorem for stationary random processes (Panchev, 1971), the mean values of these averages converge to the ensemble means in a calculable fashion. The results show that the ergodic theorem is an excellent predictor of the behavior of both the averages and the variances.

The stationarity test was applied to a number of magnetic field data sets at heliocentric distances of 1-10 AU. The length of the data records ranged from about 10 to 621 days. The degree of convergence of large-time averages to predicted values was used as a figure of merit. These analyses indicate that "good" convergence is obtained for data sets containing many ( $>10$ ) correlation times provided that significant organized structures such as isolated shocks and sector crossings are not undersampled.

These results suggest that the interplanetary magnetic field can be meaningfully viewed as a "weakly" stationary random function. Strict stationarity requires that all moments, not just means and second order correlation functions, be independent of time.

Spatial homogeneity is another possible symmetry of turbulent fluctuations (Batchelor, 1970) and is particularly important if wavenumber spectral analysis is to be performed. In the most general case, time stationarity and spatial homogeneity are separate issues. However, the solar wind is a super-Alfvénic flow which allows the approximate identification,

$$\langle B(\underline{x}, t) B(\underline{x} + \underline{r}, t) \rangle = \langle B(\underline{x}, t) B(\underline{x}, t + \tau) \rangle$$

where  $\underline{r} = -v_{sw}R$ ,  $v_{sw}$  is mean solar wind speed and  $R$  is the radial unit vector in heliocentric coordinates. This "frozen-in-flow" property is valid for phenomena occurring on a MHD time scale. If the time-stationary property holds at all points in space, the equation implies that  $B(\underline{x}, t)$  is also spatially homogeneous. The frozen-in approximation cannot be valid for spatial separations  $\underline{r}$  which approach the scale over which the bulk heliospheric quantities vary. Matthaeus and Goldstein concluded that interplanetary MHD fluctuations at frequency  $f$  correspond to spatially homogeneous structures provided that their scale  $v_{sw}/f$  is much less than the local heliocentric radial coordinate  $R$ . It thus seems reasonable to conclude that interplanetary fluctuations are both stationary and locally homogeneous with macroscopic parameters varying slowly over scale size of the heliosphere.

The equations of ideal incompressible MHD turbulence admit three known rugged invariants, the energy density,  $E = \langle B^2 + v^2 \rangle$ , the cross helicity,  $H_c = \langle \underline{v} \cdot \underline{B} \rangle$ , and the magnetic helicity,  $H_m = \langle \underline{A} \cdot \underline{B} \rangle$ . The energy density is the sum of magnetic and kinetic energy densities, with the magnetic field  $B$  written in Alfvén speed units. The magnetic helicity is a measure of departure from mirror symmetry. In the above definitions,  $B$  and  $v$  are taken to be the fluctuating field and  $A$  is assumed to be in the Coulomb gauge. In physically realistic turbulence with dissipation at the smallest spatial scales, rugged invariants are no longer constants of the motion. However, their values are not changed by the action of nonlinearities in the equations of motion. Since nonlinear couplings alone induce transfer of excitations from one length scale to another, the set of rugged invariants imposes direct restrictions on how turbulent energy transfer occurs. Matthaeus and Goldstein (1982b) developed techniques for determining the rugged invariants of three dimensional, incompressible MHD and their reduced wavenumber spectra from single spacecraft data. For this reason much of turbulence theory is couched in the vocabulary of rugged invariants and their wavenumber spectra. Here we will summarize the types of spectra most frequently seen.

The energy spectra analyzed between 1 and 10 AU show a power law wavenumber dependence of  $k^a$  with  $a = -1.55$  to  $-1.7$ . The power law region extends from scales near the correlation length, which is usually between  $10^{12}$  and  $10^{13}$  cm, down to scales at least as small as  $10^8$  cm. The power law observed near and beyond 1 AU is usually very near the Kolmogorov exponent of  $-5/3$ . The appearance of this type of power law dependence is strongly suggestive that a turbulent inertial range is being observed; this in spite of the fact that solar wind is anisotropic (Belcher and Davis, 1971) and the Kolmogorov spectral prediction presumes isotropy. These results are consistent with the hardened spectra reported inside 1 AU (Bavassano et al., 1982; Denskat and Neubauer, 1982; Woo and Armstrong, 1979).

Typical solar wind magnetic helicity spectra alternate in sign throughout the power law range and usually at lower wavenumbers as well. Because the magnetic helicity spectrum is a measure of the topological handedness of the fluctuations near a given wavenumber, the observed spectra indicate that the inertial ranges consist of both left- and right-handed magnetic structures. The net magnetic helicity for a "normal" solar wind period lies at scale lengths much larger than the magnetic correlation length  $L_c$ . This is similar to the expectation in a steady state inverse cascade (Frisch et al., 1975) or freely decaying selective decay situation (Matthaeus and Montgomery, 1980) that MHD turbulence transfers magnetic helicity preferentially to the largest allowed scales.

The cross helicity spectrum is a measure of the correlation between the velocity and magnetic fluctuations as a function of wavenumber. Similar quantities have been previously computed by Coleman (1967) and Belcher and Davis (1971) and others to describe "Alfvénic fluctuations." In highly Alfvénic intervals, the inertial range cross helicity is generally single-signed and attains a large fraction of its largest allowed values. However, it is not unusual to see the opposite sign of cross helicity at scales at or larger than the correlation length. In a wave interpretation, these wavenumber intervals correspond to inward propagation, while the inertial range has a cross helicity indicating outward propagation.

In at least one interval the observed cross helicity had both signs throughout the inertial range but with a fixed sign corresponding to outward propagation at the largest scales sampled (Matthaeus and Goldstein, 1982b). In terms of rugged invariants, some wavelengths have the opposite sign of cross helicity. These modes reintroduce nonlinearities in the dynamics and at this time we do not know if their levels are high enough to reinstate turbulent cascade processes. Recently, however, Matthaeus et al. (1983) and Matthaeus and Montgomery (1983) reported preliminary two-dimensional MHD simulation results which suggest that levels of cross helicity

even higher than those typically seen in Alfvénic periods are inadequate to prevent the development of a power law inertial range. The simulations show that the resulting nonlinear processes "almost always" act to enhance the initial alignment of magnetic and velocity fields. Even if coronal dynamics preferentially generate fluctuations correlated in the outward propagating sense, it is likely that turbulence proceeds. Even more intriguing, the simulations suggest that the sign of the cross helicity in the inertial range at one AU (always corresponding to outward propagation) may be a direct consequence of a developing turbulent cascade, and may not always be directly related to the existence of coronal Alfvén waves.

Interplanetary turbulence may be of solar origin, or may be internally generated in the interplanetary medium. In the latter case, the following dynamical scenario suggests itself. Suppose that the initially decaying turbulence is "stirred" or forced by interactions of the local fluctuations with mean field gradients and organized structures such as high speed streams (e.g., Coleman, 1968). Although the existence of an effective coupling mechanism to drive turbulence at the expense of energy in the mean fields may be viewed as conjectural, interaction regions at the leading edge of high speed streams appear to be one class of candidates for such a stirring mechanism. Other possibilities include shock waves and magnetic clouds.

Any assessment of the role of turbulence processes in the solar wind must include a determination of the time scales over which the turbulence occurs. This is well beyond our reach at this time since we cannot measure or predict forcing, dissipation or energy transfer rates without introducing unmotivated assumptions. However, a nonlinear time can be constructed by forming the ratio of the scale size of the energy containing structures to their characteristic speed. In hydrodynamics this "eddy turnover time" is an estimate of the lifetime of the system and the rate of relaxation of the excitations through cascade to smaller scales and subsequent dissipation. Using a scale of  $10^{11}$  cm and fluctuation speed of 30 km/sec gives an estimate of the MHD eddy turnover time of  $3 \times 10^4$  sec, appropriate to nominal conditions at 1 AU. Comparing this with the transit time of a 400 km/sec solar wind over an AU indicates that about 1-2 eddy turnover times are incurred for each AU of outward solar wind convection.

The fact that observed spectra of rugged invariants are qualitatively reproducible does not allow us to conclude unambiguously that turbulence is active in the solar wind. It is possible to argue that the observed power law spectra are due to noninteracting Alfvén waves. The presence of MHD structures over a wide range of spatial scale, while often associated with turbulence, is not a sensitive enough indicator to distinguish actively evolving turbulence from noninteracting waves. Continual spectra transfer from large to small scales is an essentially turbulent feature, but it is one which analysis techniques developed so far have been unable to evaluate.

Every aspect of our understanding of solar wind turbulence is primitive enough at this stage that it is fair to say that theory and observation will both have to evolve considerably before the subject is judged as complete. Problems associated with widely varying spatial scales need to be addressed. Inclusion of turbulence modeling in the large scale heliospheric expansion equations (Hundhausen, 1972; Holzer, 1979) may contribute to the resolution of heating and radial dependence problems. The degree of symmetry (isotropy, axisymmetry, etc.) of the magnetic fluctuations is only poorly known at present. At the very smallest scales the dissipation mechanism and dissipation spectra are not well understood. We anticipate that the resolution of some of these issues will play a role in our future understanding of the solar wind plasma and perhaps turbulent plasmas in general. More complete turbulence-style analysis and interpretation of solar wind observations may soon begin to shed light on the implications of turbulence for large-scale interplanetary magnetic fields (for example, through an inverse cascade or dynamo



action) and on small-scale effects such as turbulent heating and rapid reconnection (through, say, a direct cascade). Eventually it may be possible to apply the analytical machinery of turbulence modeling and closure techniques (DIA, EDQNMA, etc.) to better understand overall solar wind behavior by including turbulence effects.

## VII. EFFECTS OF WAVES AND TURBULENCE ON THE SOLAR WIND

A quarter century after Parker's original paper, the energy source of the solar wind is still not known. It is generally agreed that heating or acceleration must occur over an extended region, probably governed by MHD waves/turbulence. Indeed, several studies of radio signals passing through the corona are consistent with existence of a field of large-amplitude fluctuations extended 10-20 solar radii out from the Sun. Such a wave field would exert an outward force on the solar wind plasma, thus accelerating the wind. Dissipation of the waves would heat the plasma (note that observed interplanetary proton temperatures are much hotter than they would be if their adiabatic cooling were modified only by thermal contact with the electrons).

A number of wave-driven models of the solar wind have been developed. Models postulating a magnetoacoustic wave field invoke wave dissipation by Landau damping. Models postulating an Alfvénic wave field invoke dissipation by nonlinear processes. Both kinds of models are readily tuned to give good agreement with interplanetary observations. Certainly both processes could be acting simultaneously. Observations of the flow and fluctuations within a few tens of solar radii will be necessary to determine which, if any, wave processes are operative. Optical and radio techniques (summarized earlier in this report) are valuable probes of this near-solar region that have only begun to be applied. However, these techniques involve integration over the line of sight, and the associated ambiguities of interpretation may not admit definitive evaluation of wave-driven models. Resolution of this question probably awaits in situ measurements from a space probe in the near solar region.

The wave-driven models also are incomplete in principle. The required amplitudes are large enough that nonlinear effects may be important (indeed the dissipation of Alfvén waves would have to be nonlinear). The physics of large-amplitude hydromagnetic waves in collisionless plasma is only partly understood, and progress in this area is an important element of future progress. If the interplanetary fluctuations are in fact the remnant of fully developed turbulence in the near solar region, nonlinear processes are clearly central. It is conceivable that the wave models (based essentially on linear theory) could be completely misleading.

Interplanetary waves and turbulence are thought to play a significant role in governing a number of processes in the interplanetary medium. For example, even though the solar wind is quite collisionless, the He<sup>++</sup> generally flows faster than the H<sup>+</sup>, by as much as the local Alfvén speed. This fact is probably a signature of interaction with waves. At 0.3 AU the associated energy flux may be substantial. Thus, the helium may play an important role in the energy balance, which has not yet been elucidated. Another unexplained class of ion observations is that the ionic kinetic temperature tends to be higher than the proton kinetic temperature by a factor of order the ion mass ratio. This behavior is probably another manifestation of wave/turbulence.

Although there is convincing evidence for the solar origin of the Alfvénic component of the interplanetary fluctuations, other kinds of hydromagnetic fluctuations may be generated locally. A familiar example is the large-amplitude compressive turbulence associated with the interface

between fast and slow solar wind streams. The firehose or mirror instabilities may regulate the anisotropy of the interplanetary plasma. There is ample evidence for hydromagnetic waves generated by particle beams from planetary bow shocks. The beams in turn are eventually disrupted by the waves they have generated. Hydromagnetic waves may be significant in the acceleration and scattering of energetic charged particles by Fermi acceleration and pitch-angle scattering. Waves and turbulence may also play a large role in governing transport processes involving ions and electrons of the main solar wind plasma. These and many related topics have been discussed over the past two decades, but very little has been accomplished in the way of verifying the operation of specific mechanisms.

### VIII. REFERENCES

- Armstrong, J. W., and R. Woo, Solar wind motion within  $30 R_{\odot}$ : Spacecraft radio scintillation observations, Astron. Astrophys., **103**, 415, 1981.
- Athay R. G., The chromosphere and transition region, in *The Sun as a Star*, edited by S. Jordan, NASA SP-450, p. 85, 1981.
- Athay R. G., and O. R. White, Chromospheric and coronal heating by sound waves, Astrophys. J., **226**, 1135, 1978.
- Athay, R. G., and O. R. White, Chromospheric oscillations observed with OSO-8. IV. Power and phase spectra from C IV, Astrophys. J., **229**, 1147, 1979.
- Barnes, A., Hydromagnetic waves and turbulence in the solar wind, in *Solar System Plasma Physics, 20th Anniversary Review*, edited by C. F. Kennel, L. J. Lanzerotti, and E. N. Parker, North-Holland, Amsterdam, 1979.
- Barnes, A., Interplanetary Alfvénic fluctuations: A stochastic model, J. Geophys. Res., **86**, 7498, 1981.
- Batchelor, G. K., *The theory of homogeneous turbulence*, Cambridge University Press, New York, 1960.
- Bavassano, B., and F. Mariani, Interplanetary Alfvénic fluctuations: A statistical study of the directional variations of the magnetic field, *Proceedings of the conference Solar Wind 5* (to be published) 1983.
- Bavassano, B., M. Dobrowolny, F. Mariani, and N. F. Ness, Radial evolution of power spectra of interplanetary Alfvénic turbulence, J. Geophys. Res., **87**, 3617, 1982.
- Belcher, J. W., and L. Davis, Jr., Large-amplitude Alfvén waves in the interplanetary medium, 2, J. Geophys. Res., **76**, 3534, 1971.
- Brueckner, G. E., and J. D. F. Bartoe, Observations of high energy jets in the corona above the quiet Sun, the heating of the corona and the acceleration of the solar wind, Astrophys. J., in press 1983.
- Bruner, E. C., Jr., OSO-8 observational limits to the acoustic coronal heating mechanism, Astrophys. J., **247**, 317, 1981.

- Cheng, C. -C., G. A. Doschek, and U. Feldman, The dynamical properties of the solar corona from the intensities and line widths of EUV forbidden lines of Si VIII, Fe XI, and Fe XII., Astrophys. J., **227**, 1037, 1979.
- Christensen-Dalsgaard, J., On adiabatic non-radial oscillations with moderate or large L, Mon. Not. R. Astron. Soc., **190**, 765, 1980.
- Coleman, P. J., Wave-like phenomena in the interplanetary medium, Plan. Space Sci., **15**, 953, 1967.
- Coleman, P. J., Turbulence, viscosity and dissipation in the solar wind plasma, Astrophys. J., **153**, 371, 1968.
- Denskat, K. U., and F. M. Neubauer, Statistical properties of low-frequency magnetic field fluctuations in the solar wind from 0.29 to 1.0 AU during solar minimum conditions: HELIOS 1 and HELIOS 2, J. Geophys. Res., **87**, 2215, 1982.
- Dere, K. P., J. D. F. Bartoe, and G. E. Brueckner, HRTS observations of the quiet solar chromosphere and transition zone, Astrophys. J., submitted 1983.
- Ekers, R. D., and L. T. Little, The motion of the solar wind close to the Sun, Astron. Astrophys., **10**, 310, 1971.
- Frisch, U., A. Pouquet, J. Leorat, and A. Mazure, Possibility of an inverse cascade of magnetic helicity in magnetohydrodynamic turbulence, J. Fluid Mech., **68**, 769, 1975.
- Gough, D. O., A review of the theory of solar oscillations and its implications concerning the internal structure of the Sun, in Pulsations in Classical and Cataclysmic Variable Stars, edited by J. P. Cox and C. J. Hansen, JILA, 1982.
- Habbal, S. R., E. Leer, and T. E. Holzer, Heating of coronal loops by fast mode MHD waves, Solar Phys., **64**, 287, 1979.
- Harmon, J. K., and W. A. Coles, Spectral broadening of planetary radar signals by the solar wind, Astrophys. J., in press (1983).
- Heyvaerts, J., and E. R. Priest, Coronal heating by phase-mixed shear Alfvén waves, Astron. Astrophys., **117**, 220, 1983.
- Hollweg, J. V., Some physical processes in the solar wind, Rev. Geophys. Space Phys., **16**, 689, 1978.
- Hollweg, J. V., Surface waves on solar wind tangential discontinuities, J. Geophys. Res., **87**, 8065, 1982.
- Hollweg, J. V., Resonances of coronal loops, Solar Phys., in press 1983.
- Hollweg, J. V., M. K. Bird, H. Holland, P. Edenhofer, C. T. Stelzried, and B. L. Seidel, Possible evidence for coronal Alfvén waves, J. Geophys. Res., **87**, 1, 1982.

- Holzer, T. E., The solar wind and related astrophysical phenomena, in *Solar System Plasma Physics*, ed. C. F. Kennel, L. J. Lanzerotti and E. N. Parker, vol. 1, North-Holland, 1979.
- Hundhausen, A. J., *Coronal expansion and solar wind*, Springer-Verlag, 1972.
- Ionson, J. A., Resonant absorption of Alfvénic surface waves and the heating of solar coronal loops, *Astrophys. J.*, **226**, 650, 1978.
- Ionson, J. A., Resonant electrodynamic heating of stellar coronal loops: an LRC circuit analog, *Astrophys. J.*, **254**, 318, 1982.
- Kohl, J. L., H. Weiser, G. L. Withbroe, W. H. Parkinson, E. M. Reeves, R. H. Munro, and R. M. MacQueen, Measurements of coronal kinetic temperatures from 1.5 to 3 solar radii, *Astrophys. J.*, **241**, L117, 1980.
- Kohl, J. L., G. L. Withbroe, C. A. Zapata, and G. C. Noci, Spectroscopic measurements of solar wind generation, in *Solar Wind 5*, edited by M. Neugebauer, NASA in press 1983.
- Leibacher, J., and R. F. Stein, Oscillations and pulsations, in *The Sun as a Star*, edited by S. Jordan, NASA SP-450, p. 263, 1981.
- Liebenberg, D. H., R. J. Bessey, and B. Watson, Coronal emission line profile observations at total solar eclipses, *Solar Phys.*, **44**, 345, 1975.
- Matthaeus, W. H., and M. L. Goldstein, Stationarity of magnetohydrodynamic fluctuations in the solar wind, *J. Geophys. Res.*, **87**, 10347, 1982a.
- Matthaeus, W. H., and M. L. Goldstein, Measurement of the rugged invariants of magnetohydrodynamic turbulence in the solar wind, *J. Geophys. Res.*, **87**, 6011, 1982b.
- Matthaeus, W. H., and D. C. Montgomery, Selective decay hypothesis at high mechanical and magnetic Reynolds numbers, *Ann. N.Y. Acad. Sci.*, **357**, 203, 1980.
- Matthaeus, W. H., M. L. Goldstein, and D. C. Montgomery, Turbulent generation of outward traveling interplanetary Alfvénic fluctuations, *Phys. Rev. Lett.*, in press, 1983.
- Matthaeus, W. H., and D. C. Montgomery, Dynamic alignment and selective decay in MHD, *Statistical Physics and Chaos in Fusion Plasma*, C. W. Horton, Jr. and L. E. Reichl eds. John Wiley, in press, 1983.
- Panchev, S., *Random functions and turbulence*, Pergamon Press, 1971.
- Parker, E. N., The rapid dissipation of magnetic fields in highly conducting fluids, *Geophys. Astrophys. Fluid Dyn.*, **22**, 195, 1982.
- Roberts, B., and A. Mangeney, Solitons in solar magnetic flux tubes, *Mon. Not. R. Astron. Soc.*, **198**, 7P, 1982.
- Roberts, B., et al., submitted to *Nature* 1983.
- Scott, S. L., Density spectrum and velocity of the solar wind inferred from scintillation observations, PhD thesis, Univ. of Calif., San Diego, La Jolla, 1978.

- Spruit, H. C., Theory of photospheric magnetic fields, in proc. IAU symposium 102 "Solar and Stellar Magnetic Fields," in press 1982.
- Stein, R. F., Stellar chromospheric and coronal heating by magnetohydrodynamic waves, Astrophys. J., **246**, 966, 1981.
- Stein, R. F., and J. Leibacher, Waves in the solar atmosphere, Ann. Rev. Astron. Astrophys., **12**, 407, 1974.
- Thomas, J. H., Magneto-atmospheric waves, Ann. Rev. Fluid Mech., **15**, 321, 1983.
- Withbroe, G. L., J. L. Kohl, H. Weiser, G. C. Noci, and R. H. Munro, Analysis of coronal H I Lyman alpha measurements from a rocket flight on 1979 April 13, Astrophys. J., **254**, 361, 1982a.
- Withbroe, G. L., J. L. Kohl, H. Weiser, and R. H. Munro, Probing the solar wind acceleration region using spectroscopic techniques, Space Sci. Rev., **33**, 17, 1982b.
- Woo, R., and J. W. Armstrong, Spacecraft radio scattering observations of the power spectrum of electron density fluctuations in the solar wind, J. Geophys. Res., **84**, 7288, 1979.
- Zugzda, Y. D., and V. Locans, Tunneling and interference of Alfvén waves, Solar Phys., **76**, 77, 1982.



## CHAPTER 5

# COUPLING OF THE SOLAR WIND TO THE MAGNETOSPHERE

### WORKING GROUP MEMBERS

J. K. Alexander, Chairman  
*NASA/Goddard Space Flight Center*

L. F. Bargatze  
*University of California, Los Angeles*

J. L. Burch  
*Southwest Research Institute*

T. E. Eastman  
*University of Iowa*

J. G. Lyon  
*Naval Research Laboratory*

J. D. Scudder  
*NASA/Goddard Space Flight Center*

T. W. Speiser  
*University of Colorado*

G. H. Voigt  
*Rice University*

C. -C. Wu  
*University of California, Los Angeles*

**CHAPTER 5**  
**COUPLING OF THE SOLAR WIND TO THE MAGNETOSPHERE**

I.	Introduction	5-3
II.	Major Elements of Solar Wind-Magnetosphere Coupling	5-6
III.	Solar Wind Coupling Mechanisms	5-11
IV.	Global Modeling	5-18
V.	Summary and Outstanding Problems	5-20
VI.	References	5-22



## I. INTRODUCTION

The interaction between the solar wind and the Earth's magnetosphere is particularly interesting because the geomagnetic field is strong enough to present a significant obstacle to the solar wind flow but weak enough to allow the magnetosphere to respond sensitively to changes in solar wind plasma and magnetic field parameters. Simple pressure balance arguments long ago predicted the existence of the magnetosphere as a cavity in the solar wind within which plasma motions are controlled by the geomagnetic field and from which the solar wind plasma is almost totally excluded. In fact, the interaction is much more complex than at first supposed. As yet unexplained features of the coupling process allow a small, but significant, fraction of the solar wind plasma flux to enter the magnetosphere. In addition, the momentum carried by the solar wind is transferred efficiently to the outer regions of the magnetosphere. The coupling between the solar wind and the magnetosphere is, in fact, so effective that it controls almost totally the dynamic configuration of the magnetosphere, the various plasma populations that exist within it, and the plasma disturbances that map magnetically into the high latitude ionosphere. The large body of knowledge that now exists on magnetospheric structure and plasma populations therefore places a number of constraints on, and provides a number of clues to, the nature of the coupling process.

We begin in this section with a brief description of the configuration of the magnetosphere and its various plasma regimes and boundaries. Next, a set of four basic questions regarding our understanding of the interaction and our ability to model it are posed. These four questions serve as an outline for the remainder of the chapter, in which our present knowledge and prospects for future progress are examined in some detail.

### MAGNETOSPHERIC PLASMA REGIMES AND BOUNDARIES

The various magnetospheric plasma domains and their boundaries are illustrated in Figure 5-1. On the sunward side of the magnetosphere, a collisionless bowshock acts to deflect the oncoming solar wind flow via electric fields set up within the shock. The solar wind plasma is partially thermalized and slowed down at this shock. Most of this shocked solar wind or magnetosheath plasma flows around the magnetosphere and continues downstream. However, a small fraction of the magnetosheath plasma interacts with the outer boundary of the magnetosphere, the magnetopause. Within the magnetosphere, charged particle motion is dominated by forces due to the Earth's magnetic and electric fields. As a result of the plasma-magnetic field interactions near the magnetopause, the magnetospheric boundary layer is generated on field lines located earthward of the magnetopause, although the plasma distributions sampled there are closely related to those observed in the nearby magnetosheath.

The major low altitude boundary is the ionosphere, which constitutes a conducting surface linked to the outer magnetosphere via highly conducting magnetic "field lines". One must be cautious about assuming that magnetic field lines are always equipotentials, however. Increasingly evident is the fact that parallel (to B) electric fields are often important in magnetospheric processes, even when collisional effects are negligible. Ohmic heating and other dissipative effects in the ionosphere can also act as a "load" for current generation within the outer magnetosphere. A second major boundary of the inner magnetosphere is the plasmapause, which bounds the inner region of corotating plasma, the plasmasphere. It also constitutes an approximate inner bound for the ring current, which is controlled by gradient and curvature drifts, and at times, can overlap the inner edge of the plasma sheet.

The Earth's geomagnetic tail extends downstream to at least several times lunar distance due to the balance between plasma pressure gradients, magnetic stresses and streaming velocities.

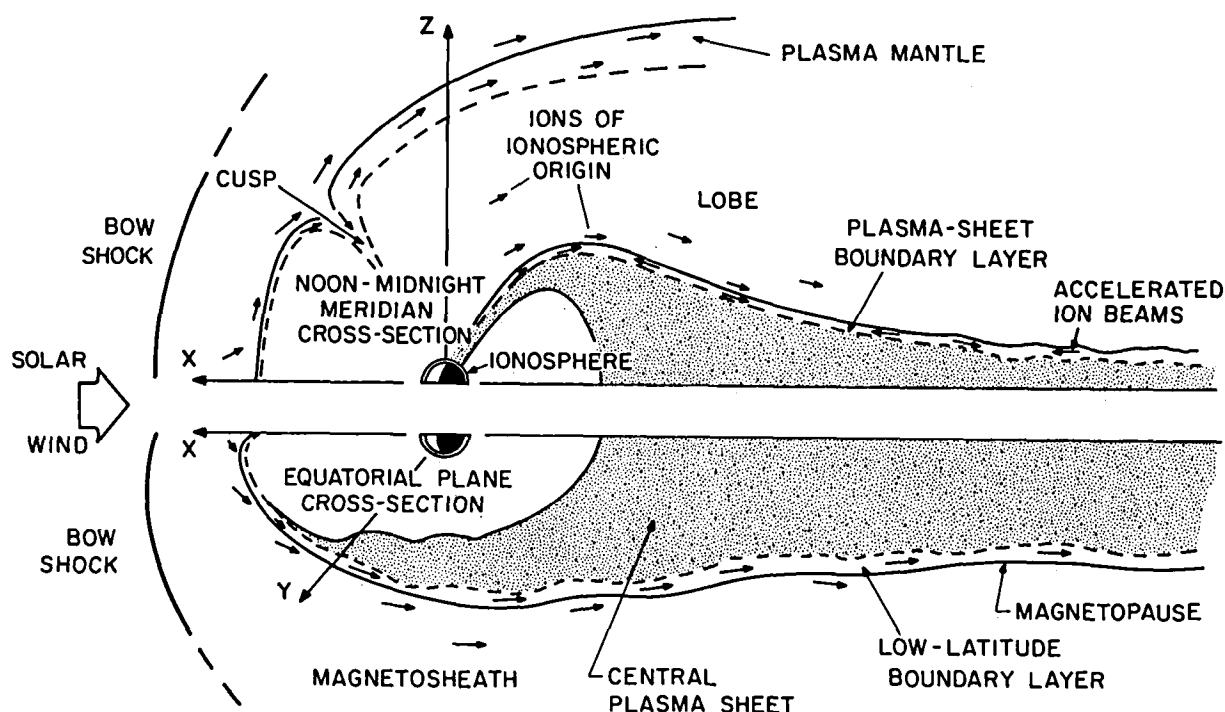


Figure 5-1. A schematic view of plasma transport in magnetospheric boundary layers. The upper half of the diagram is a cross-section in the noon-midnight meridian plane; the lower half represents an equatorial plane cross-section.

Within this magnetotail, a large region of non-rotating hot plasma, the plasma sheet, is formed. As shown in Figure 5-2, the plasma sheet is bounded in both hemispheres by the extended polar cap or magnetotail lobe regions. Whereas the magnetosheath has a high plasma  $\beta$  (the ratio of plasma to magnetic field energy density) and the plasma sheet typically has a medium  $\beta$  ( $\sim 1$ ) plasma, the lobe region usually has very low  $\beta$ . Within the lobe regions, cold ion streams flowing away from the Earth are commonly observed. These ion streams appear to be of ionospheric origin as they contain significant fluxes of  $O^+$  and  $He^+$ .

A plasma sheet boundary layer has recently been identified at the interface between the tail lobes and the plasma sheet. This region is characterized by highly anisotropic ion velocity distributions including energetic ion beams and, concurrently, the same cold, low-energy ion streams that occur in the adjoining lobe region. The energetic ion beams are magnetically mirrored near the Earth, resulting in counter-streaming ion beams throughout the plasma sheet boundary layer.

High speed plasma flows observed within the magnetosphere are generally limited to the boundary layers, which act as the primary transport regions of the Earth's magnetosphere. Figure 5-3 illustrates the relationship of the magnetospheric regimes and boundaries described above. The inner ("ionosphere") and outer ("solar wind") boundaries act as source regions, whereas the plasma sheet and lobes act as storage regions primarily for plasma and field energy, respectively. Most of the plasma transport within the magnetosphere as well as the field-aligned current systems providing the ionosphere-magnetosphere interaction occurs within the boundary layers. These boundary layers in turn are the locus of the major acceleration regions of the magnetosphere including the upward acceleration of ions at low altitude along auroral field lines and the earthward acceleration of energetic ion beams.

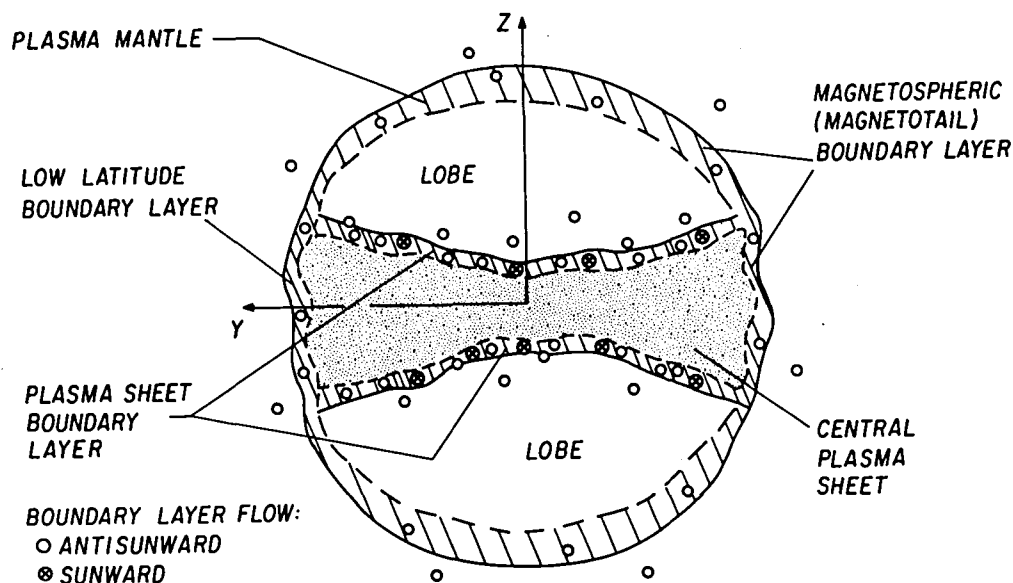


Figure 5-2. A magnetotail cross-section sketch of the major magnetospheric plasma regimes. This view is orthogonal to the cross-sections shown in Figure 5-1.

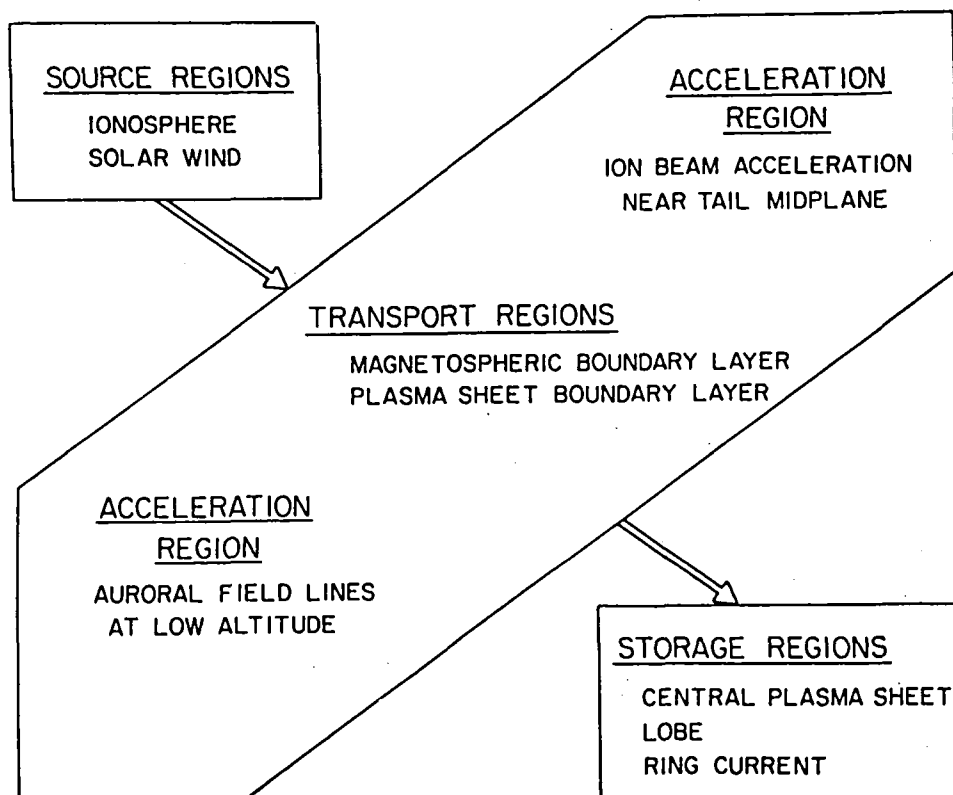


Figure 5-3. A schematic description of the origin, transport, acceleration and storage of magnetospheric plasma. Notice the importance of boundary layers as major transport paths.

## MAJOR QUESTIONS CONCERNING SOLAR WIND–MAGNETOSPHERE COUPLING

We now have a general descriptive knowledge of the magnetosphere and its responses to changes in solar wind parameters. The major outstanding questions concern the nature of the coupling processes themselves and the global magnetospheric dynamics that would theoretically result from these processes. The importance of the latter element lies in the fact that many of our observations are not of the actual coupling process but of the dynamical behavior that results from it.

These major issues are posed in the following four questions:

1. What processes are responsible for the entry of solar wind plasma into the magnetosphere, and how do they respond to changes in solar wind parameters?
2. What processes are responsible for the transfer of solar wind energy and momentum into the magnetosphere? What are the resulting boundary conditions on the magnetopause, and how do they respond to changes in solar wind parameters?
3. What is the role of magnetospheric boundary layers in the dissipation and generation of electrical energy and as plasma transport regions?
4. What global MHD (magnetohydrodynamic) effects are produced by various coupling processes as functions of time-varying, solar wind parameters?

In the remainder of this chapter, these four questions are addressed as we describe the present state of knowledge on solar wind-magnetosphere coupling and attempt to chart a course for future progress in this field.

## II. MAJOR ELEMENTS OF SOLAR WIND–MAGNETOSPHERE COUPLING

In many ways the magnetized solar wind plasma displays the fluid-like behavior predicted by hydromagnetic theory in its interaction with the geomagnetic field. A thin (~1000 km) bowshock impedes the solar wind flow and thermalizes it to produce a subsonic flow in the magnetosheath. A large-scale current system confines the magnetosphere to a cavity which, on the day-side, has very nearly the shape predicted by gas dynamic theory for flow around a blunt obstacle. However, the greatly extended geomagnetic tail, the continuous presence of large-scale plasma convection within the magnetosphere, the large-scale contractions of the dayside magnetosphere, and the impulsive nature of magnetospheric substorms require the hydromagnetic approximation to break down along the magnetopause and in the magnetotail. Although we have an impressive set of evidence from macroscopic observables that non-MHD plasma processes control the energetics of the solar wind-magnetosphere interaction with some predictability, only scant data exist on the nature of the processes themselves. This situation prevails because the candidate processes are apparently quite limited in spatial extent with plasma effects that may also be very localized in phase space. Therefore, experimental resolution of the coupling question requires a diverse program of both in situ and remote sensing observations of phenomena throughout the magnetosphere and in the nearby solar wind.

In the following paragraphs we review seven important elements that must be accommodated by a successful model of solar wind-magnetosphere coupling. In each case a preliminary assessment of the ability of contemporary models to explain the observation is made.

## THE ENTRY EFFICIENCY OF SOLAR WIND PLASMA

Only a very small fraction ( $\sim 0.5\%$ ) of the solar wind plasma flux impinged on the magnetospheric surface is captured by the magnetosphere. This low entry efficiency places no demanding requirements on models of solar wind-magnetosphere coupling. It does mean, however, that any efficient plasma coupling processes must be very localized in space or time. Indeed, we know that magnetosheath plasma is injected directly into the cusps and is transported to higher latitudes by strong convection electric fields. A large fraction of the injected cusp plasma mirrors above the atmosphere and flows upward to populate the plasma mantle (see Figure 5-1). The cusp injection region is very narrow (a few gyroradii) in latitude and may result from merging between geomagnetic and magnetosheath magnetic fields. Diffusion of cusp plasma toward the equator may also occur and contribute to the low-latitude boundary layer. Gradient and curvature drift and cross-field diffusion driven by wave turbulence should also transport some magnetosheath plasma across the low latitude magnetopause and into the boundary layer. However, as presently envisioned, these processes are not efficient enough to populate the low-latitude boundary layer by themselves.

One localized, highly efficient, low-latitude, plasma entry process--impulsive plasma injection--has been proposed. In this model magnetosheath density enhancements "burrow" their way through the magnetopause by first "denting" it and then gaining full access by the action of diamagnetic surface currents. Although this model may provide a source of flowing plasma to the boundary layer, we note below that magnetic merging may need to occur in association with this process before it can become a significant source of magnetospheric energy or momentum.

The primary reservoirs of magnetospheric plasma are the plasma sheet and plasmasphere. For the plasma sheet, it is estimated that a source capable of supplying about  $10^{26}$  ions  $\text{sec}^{-1}$  must exist to maintain it. However, populating the plasma sheet is at least a two-step process. First, magnetosheath plasma must gain entry to the low-latitude boundary layer or the plasma mantle with negligible change in temperature, but with reduced densities and flow velocities. Next, the plasma must be transported, or injected, into the plasma sheet and be energized to average energies of a few keV.

## RECONNECTION OF GEOMAGNETIC AND INTERPLANETARY MAGNETIC FIELDS

Some significant fraction of the polar caps is magnetically connected to the solar wind. This fraction grows significantly during periods when the interplanetary magnetic field (IMF) has strong southward components and shrinks drastically during extended periods of northward IMF. The evidence for this open magnetic topology, as first proposed by Dungey (1961), has come from extensive studies of the entry of energetic solar flare particles into the magnetosphere, from the observation of asymmetries in east-west convection and Birkeland currents in the dayside magnetosphere, and most notably from the strong correlation of these asymmetries with the azimuthal direction of the IMF. The evidence for an open magnetosphere is overwhelming. However, the boundary between the open and closed regions of the magnetosphere does not maintain a fixed relationship to other magnetospheric regions. Thus, at any given universal time or local time, some or all of the auroral oval and even significant portions of the polar caps may be threaded by closed field lines. As the unambiguous local identification of magnetic topology (whether open or closed) is not yet possible, there continue to be many unanswered questions on this topic.

Because plasmas are observed to flow across the topological boundaries separating closed and open magnetospheric regions, some process of magnetic merging or reconnection must take place (Vasyliunas, 1975). The fact that these topological boundaries are also sites of strong Birkeland

current flow, plasma acceleration, and convection reversals suggests that the process responsible for the reconnection is also a significant factor in the transfer of solar wind energy and momentum to the magnetosphere.

Sometimes the term "interconnection" has been used in discussing the open magnetospheric topology. Reconnection specifically implies an electric field,  $E_{\parallel}$ , along a magnetic separator, the line of intersection of two separatrix surfaces dividing space into different magnetic cells. (See Chapter 1, Reconnection of Magnetic Fields.) As interconnection implies to some an open topology without reconnection, the implications should be discussed. An open topology without reconnection would imply zero electric field along the separator. This would imply that magnetic flux bounded by the separatrix surfaces cannot change, i.e., the open topology would have to be primordial. Furthermore, since the field line ends that are located far away in the solar wind have an electric field between them and near the separator they cannot support such a field, there must be large parallel electric fields along such solar wind field lines when topological flux changes are occurring. Although certainly not impossible, the effects of such strong parallel electric fields should be observable.

### CONTROL OF DAYSIDE CONVECTION PATTERNS BY THE EAST–WEST IMF COMPONENT

Ground-based magnetometer measurements of ionospheric currents and spacecraft observations of plasma convection and Birkeland currents have established a direct relationship between the east-west component of the IMF and dayside convection patterns (the Svalgaard-Mansurov effect). In the Northern Hemisphere, eastward IMF components (positive  $B_y$  produce mostly westward convection in the polar cusp and a concentration of convection on the dawn side of the polar cap. Conversely, westward IMF components produce eastward convection in the cusp and enhanced convection in the dusk sector of the polar cap. These effects, which are reversed in the Southern Hemisphere, are almost universally attributed to the interconnection of geomagnetic and magnetosheath magnetic fields in the quadrant of the dayside magnetosphere where they are most nearly antiparallel. The magnetic tension caused by the diversion of magnetosheath plasma around the magnetosphere is thought to produce strong plasma flows across the cusp and away from the quadrant of magnetic interconnection.

The control of dayside convection asymmetries by the IMF is one of the strongest pieces of evidence for an open magnetosphere and is one of the key phenomena that must be accounted for by models of solar wind-magnetosphere coupling. However, the likelihood that a significant portion of the low-latitude boundary layer, where antisunward plasma flow is observed, has a closed magnetic topology must also be explained.

### MAGNETIC TOPOLOGY OF THE LOW–ALTITUDE BOUNDARY LAYER

Observations in the dayside low-latitude boundary layer have shown that in general the plasma flow velocities are antisunward and somewhat lower than those observed in the adjacent regions of the magnetosheath. Trapped energetic electron distributions observed in these same regions suggest (but do not prove) that a significant portion of the boundary layer is threaded by closed magnetic field lines. However, recent studies have shown that electron "trapping" can occur on open field lines (in the minimum B region of the magnetopause). This question of magnetic topology needs to be sorted out in future studies. As pointed out earlier, impulsive plasma injection is the only proposed process that predicts the localized efficient injection of magnetosheath plasma into a boundary layer that is on closed field lines. Impulsive injection is, however, at present only a conceptual model that has not been carried to the point of making quantitative predictions.

It is possible that electromagnetic instabilities produced by velocity shears at the magnetopause produce kinetic Alfvén waves and turbulent diffusion, which may contribute to the population of the boundary layer with magnetosheath plasma. There is ample evidence for the existence of such Kelvin-Helmholtz type instabilities both at the magnetopause and near the inner surface of the low-latitude boundary layer.

## CORRELATION OF GEOMAGNETIC ACTIVITY WITH SOUTHWARD IMF COMPONENTS

When the IMF takes on a persistent southward component, gross changes in the configuration of the magnetosphere occur, and substorm activity increases markedly. The changes in magnetospheric configuration can be understood in terms of an initially unbalanced intensification of antisunward convection that is later balanced by an enhanced return flow within the magnetosphere. Compared to quiet times this new configuration is characterized by a net transfer of magnetic flux from the dayside to the magnetotail that is evidenced by an earthward displacement of the magnetopause, an equatorward migration of the polar cusp, and an increase in the size of the auroral oval. The substorm activity that normally occurs during such periods makes these effects more prominent.

A variety of analysis techniques have been utilized to determine the statistical nature of substorms. Cross-correlation studies between interplanetary parameters and geomagnetic indices have established the relative importance of the solar wind flow speed,  $V$ , the IMF magnitude and especially the IMF  $B_z$  component. Most recently, Baker et al. (1983) correlated various solar wind inputs with the AE index and  $U_T$  ( $U_T$  is approximately proportional to the total rate of energy dissipation within the magnetosphere). They find a peak correlation when AE or  $U_T$  lag the solar wind input by about 40 minutes.

Good correlations between geomagnetic indices and  $VB_s$  and  $V^2B_s$  (where  $B_s$  is the southward component of the IMF) suggested that the magnetosphere behaves like a half-wave rectifier of the solar wind motional electric field. In contrast, it has been proposed that a parameter more closely related to geomagnetic activity is  $\epsilon = VB^2 1_0^2 \sin^4(\theta/2)$  where  $\theta = \tan^{-1} |B_y/B_z|$  for  $B_z \geq 0$  or  $\theta = \pi - \tan^{-1} |B_y/B_z|$  for  $B_z < 0$  and  $1_0$  = linear scale ( $\sim 7R_E$ ) of the solar wind-magnetosphere coupling region (Perreault and Akasofu, 1978). These authors identify  $\epsilon$  as the dynamo power delivered from the solar wind to the open magnetosphere, and they find that is roughly proportional to the AE index,  $\epsilon \sim \text{AE} \times 10^{16} \text{ erg sec}^{-1}$ .

Although no one doubts the strong correlation of southward IMF components with substorm activity, the difficulty in accurately timing the onset time of the substorm expansion phase has resulted in a cause-effect controversy. Two distinct temporal models of solar wind-magnetosphere interaction have been presented. In the driven model, the magnetosphere responds directly to variations in solar wind input. Conversely, in the energy storage-release model, the magnetosphere responds to the solar wind by storing a large fraction of coupled energy in the magnetotail as reconnected magnetic flux is convected tailward. Later, the stored energy is impulsively released from the tail at the onset of a substorm expansion phase when the magnetotail configuration becomes unstable.

The driven and storage-release models have been compared using linear prediction analysis (Bargatze et al., 1983). The magnetosphere apparently responds to solar wind input ( $VB_s$  to AL index) over a 20-minute time scale and a 60-minute time scale. These response time scales suggest that both solar wind driven phenomena and magnetotail energy release phenomena are important facets of magnetospheric substorm. (See also Chapters 7 and 8).

Reconnection provides the most convenient coupling model for explaining the gross features of the magnetospheric response to southward IMF components. Estimates of the efficiency with which the solar wind electric field is transmitted into the magnetosphere range from a few percent up to 20% for strong southward IMF components. This entry efficiency is significantly higher than that of solar wind plasma entry noted above in the section entitled the "Entry Efficiency of Solar Wind Plasma." In fact, coupling of the solar wind electric field into the magnetosphere is apparently direct enough that even low-latitude phenomena such as the ring current are found to respond sensitively to its variations.

### SUNWARD CONVECTION IN THE POLAR CAPS

Normally both polar caps map magnetically into the deep magnetotail and ultimately into the solar wind. As expected, antisunward plasma convection is observed throughout these regions. However, in exceptionally quiet times when the IMF has a strong northward component, sunward convection is observed over a significant portion of the polar caps. These cells of sunward convection appear to be associated with Sun-aligned auroral arcs and with localized regions of closed magnetic topology. How these regions map into the magnetotail is a matter of some conjecture. Although no quantitative, global model of solar wind-magnetosphere coupling for northward IMF has yet been proposed, the existence of closed flux tubes and sunward convection for these conditions was predicted by Dungey's original merging theory over two decades ago.

### FLUX TRANSFER EVENTS AT THE DAYSIDE MAGNETOPAUSE

Ultimately, successful models of the solar wind-magnetosphere interaction must be based on detailed observations of plasma interactions at the magnetopause. Paschmann et al. (1979), Sonnerup et al. (1981) and Gosling et al. (1982) have reported case studies of 12 events near the nose of the magnetopause in the framework of the quasi-steady reconnection picture of Levy et al. (1964). These events are consistent with tangential stress balance at a rotational discontinuity in an anisotropic plasma and appear to meet the necessary conditions for steady reconnection. Aggson et al. (1983) have recently reported electric field measurements that confirm the rotational characteristics of the boundary and its time-steady character.

The ISEE spacecraft have also observed phenomena consistent with the sporadic occurrence of low latitude merging during periods of southward IMF. These phenomena, referred to as flux transfer events (Russell and Elphic, 1979), have been interpreted as newly reconnected magnetic flux tubes that are being pulled away from the subsolar region by the magnetosheath flow. These flux tubes contain energetic particle fluxes and magnetic field strengths that are above magnetosheath levels and more typical of the magnetosphere. The occurrence frequency of flux transfer events has been estimated to be at least 0.4 per hour which, although perhaps coincidentally, is on the order of the substorm occurrence frequency.

In reviewing the major elements of solar wind-magnetosphere coupling it is evident that merging or reconnection models are conceptually able to explain many of the phenomena occurring in the cusp and polar cap. However, the low-latitude magnetopause and boundary layer phenomena may require consideration of closed-model processes. It may therefore be unnecessary for a single process to explain solar wind plasma entry and energy and momentum transfer over the entire magnetopause. A unified theory must, therefore, consider both closed and open model processes, be three-dimensional and global, and be able to follow variations in solar wind plasma and magnetic field parameters.



### III. SOLAR WIND COUPLING MECHANISMS

Several quite different mechanisms have been proposed for the transfer of solar wind plasma, energy and momentum into the magnetosphere. No one of these processes can by itself explain even qualitatively the major features of the interaction which were summarized in Section II. Some combination of closed model and open model processes may, in fact, be appropriate.

Some of the proposed coupling processes, such as simple neutral point plasma injection near the cusp and diffusive plasma entry across closed magnetic field lines, may allow solar wind plasma to enter the magnetosphere but may not contribute significantly to the energization or convection of magnetospheric plasma. Other processes, such as reconnection, may primarily act to couple the solar wind electric field (i.e., momentum) into the magnetosphere, with plasma entry as a secondary effect. Moreover, while transfer of momentum to magnetospheric plasma may be initiated by processes such as reconnection on the dayside, the energy flow (as represented by the electromagnetic Poynting flux) is directed primarily into the magnetotail. For these reasons one must evaluate separately the importance of the candidate coupling processes as they contribute to plasma entry, momentum transfer and energy transfer.

#### PROPOSED COUPLING PROCESSES

The processes that have been proposed to explain the major features of solar wind-magnetosphere coupling fall into two categories: open model processes involving magnetic reconnection and closed model processes which involve plasma transport across magnetic field lines. The first category includes steady-state reconnection (Petschek, 1964) and localized sporadic reconnection as possibly evidenced by flux transfer events (Russell and Elphic, 1979). The second category includes the general class of diffusive plasma entry processes and the recently proposed impulsive plasma injection model (Lemaire, 1977).

Discussion of these four processes appear in the following paragraphs.

1. Diffusive Entry. Because plasma in the magnetosheath as well as in the magnetosphere is collisionless, diffusive processes are commonly described in terms of anomalous resistivity. Explanations for anomalous resistivity include wave-particle interactions and instabilities associated with the existence of the magnetopause discontinuity. A purely dissipative process can be active even in the context of a "closed" magnetosphere model; i.e., a model in which the magnetopause is assumed to be a tangential discontinuity.

Kinetic instabilities also can provide diffusive transport in the solar wind-magnetosphere interaction problem. Eviatar and Wolf (1968) first examined the effect of pitch-angle scattering near the magnetopause based upon field fluctuations. More recently, Tsurutani and Thorne (1982) have used measured plasma wave spectra and a cyclotron-resonant scattering mechanism to explain the observed mass and momentum transport of electrons across the magnetopause.

Other microinstabilities have been proposed to generate waves capable of scattering particles across the magnetopause to form a boundary layer. Some of these are: the kinetic Alfvén wave instability (Hasegawa and Mima, 1978); the lower-hybrid-drift instability (Gary and Eastman, 1979); and the electrostatic flow shear instability (Gary and Schwarz, 1980).

Haerendel (1978) proposed an eddy (microscopic) interchange transport process at the magnetopause as an entrance mechanism for magnetosheath plasma. This model, however, needs to be developed quantitatively in order to estimate its importance.

2. Steady-State Reconnection. The process of "magnetic reconnection" or "merging" is associated with a violation of the frozen flux condition

$$\underline{E} + (\underline{v} \times \underline{B}) = 0 \quad (1)$$

at the magnetopause in the associated diffusion region. In this region, field energy is converted to plasma energy. A steady-state picture would result in a topology of field lines connected through the boundary as illustrated in Figure 5-4. Magnetosheath particles can penetrate the magnetopause via simple guiding center motion along open field lines. In the reconnection model, magnetosheath plasma has direct access to the polar cusp along open field lines. This access combined with antisunward convection provides a natural explanation of the plasma mantle (Figure 5-1).

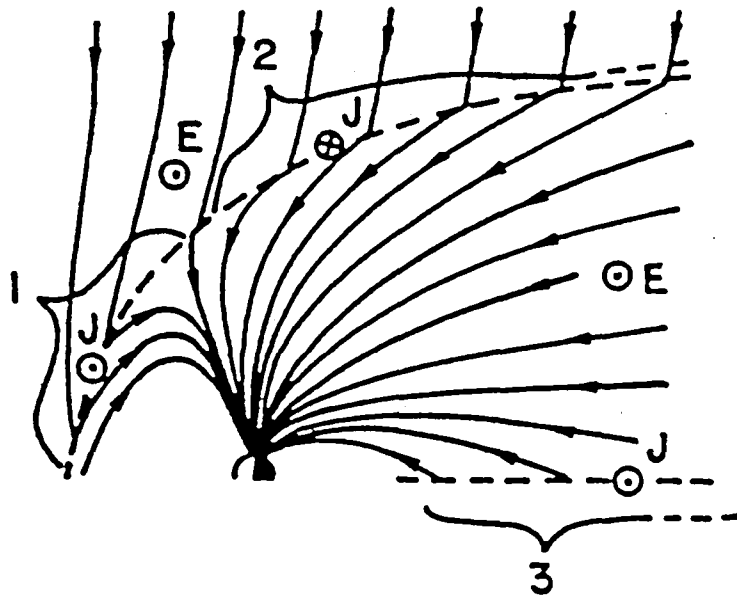


Figure 5-4. A simple view of an open magnetosphere illustrating the magnetopause (1 and 2) and magnetotail (3) current sheets. In region 3,  $\underline{j} \cdot \underline{E}$  is positive whereas in region 2,  $\underline{j} \cdot \underline{E}$  is negative. The sign of  $\underline{j} \cdot \underline{E}$  in region 1 depends on the importance of open vs. closed model processes. (Adapted from Hill, 1983.)

3. Impulsive Plasma Penetration. A potentially efficient localized process of plasma injection into the magnetospheric boundary layer is that of impulsive penetration of finite magnetosheath plasma filaments (Lemaire and Roth, 1978). Schindler (1979) pointed out that impulsive penetration of plasma filaments can occur with neither the necessity of a generally diffusive magnetopause nor the classical reconnection pattern. Although this mechanism seems to predict an increased correlation between solar wind fluctuations and geomagnetic activity, whether or not impulsive penetration is an important process for plasma and energy transfer into the magnetosphere is at present an unresolved question. Schindler noted that for infinite filaments the idealized MHD assumption provides for entry only under highly restricted conditions (equal penetration only for parallel and antiparallel magnetic field geometries). This restriction suggests that resistive effects must be significant for such impulsive injection processes. Such resistive processes are part of related work on "patchy" or sporadic reconnection as well as finite plasma filaments interacting with the magnetopause with dissipation in the ionosphere or other external "loads".

4. Flux Transfer Events--Localized Patchy Reconnection. Flux transfer events (FTEs), originally identified in the magnetic field data by Russell and Elphic (1978), have recently been reexamined in light of the plasma observations (Paschmann et al., 1982 and Scudder et al., 1983) to ascertain their dynamical signatures and their importance for solar wind-magnetosphere coupling. These events have been interpreted by Russell and Elphic as signatures of patchy reconnection at sites distant from the spacecraft. Apparent kinks in the magnetic field are suggestive of a pattern of reconnected flux tubes moving essentially tangential to the nominal magnetopause which is not moving significantly in the radial direction. They have characteristic spatial dimensions of  $1-2 R_E$  as intercepted and are usually, but not always, accompanied by electron heat flux layers (Scudder et al., 1983). As illustrated schematically in Figure 5-5, these heat flux layers frame expansion fan layers which represent signatures of plasma acceleration (+ or -) consistent with a partial traversal of a rotational discontinuity.

The electron pressure anisotropies are strongly enhanced due to the admixture of cold field-aligned plasmas similar to some measurements made in the mid-altitude polar cusp (Burch et al., 1983) suggesting that FTEs are topologically open flux tubes even down to the scale of 10 eV electrons (Scudder et al., 1983). Earlier the mapping done by Speiser et al. (1981) and Speiser and Williams (1982) suggested from arguments based on particle flux anisotropies at  $>10$  keV energy that these topologies are open. The identification of these heat flux layers with the magnetic separatrix (see Figure 5-5) supports the solar wind-magnetosphere coupling roles of FTEs because the dynamics organized within this separatrix are clearly ordered as expected for spatially limited acceleration attendant to reconnection. Moreover, the existence of a separatrix is necessary in order that FTEs be associated with reconnection and solar wind-magnetosphere coupling. Since the separatrix and signatures of moving plasma across this surface appear to exist, we take this as strong evidence for local reconnection as defined by Vasyliunas (1975). Scudder et al. (1983) have estimated that FTEs can supply the magnetic flux to the tail as required by observed substorm time scales.

## MOMENTUM AND ENERGY TRANSFER

In principle, the concept of the open magnetosphere allows us to map the solar wind  $\mathbf{v} \times \mathbf{B}$  electric field down into the ionosphere along open, interconnected magnetic field lines. In this respect, the polar cap potential distribution can be regarded as an image of the plasma flow and magnetic field conditions in the solar wind. During conditions of southward IMF, a polar cap potential drop of about 50 kV is typically observed. Maximum observed potential drops are of order 200 kV, and a good correlation is found with the magnitude of the southward IMF. A particular consequence of the solar wind electric field mapping is the explanation of the Svalgaard-Mansurov effect. This effect explains an asymmetry in the polar cap electric field as a consequence of the east-west component of the interplanetary magnetic field.

The knowledge of the function  $B_n(x)$ , the magnetic field normal component along the magnetopause, is crucial since it determines the "openness" of the magnetosphere and, hence, the size of the polar oval and the ionospheric convection  $\mathbf{E}$ -field pattern.

Although there is a general agreement that the magnetosphere is open, one can consider transfer of energy and momentum across a closed magnetopause as well. This process is associated with a "viscous-like" interaction as originally proposed by Axford and Hines (1961). That at least some minimal viscous-like interaction is appropriate is supported by a study by Reiff et al. (1981). These authors found that the polar cap potential never dropped below  $\sim 35$  kV, even for northward IMF. Wygant et al. (1983) have revised this figure downward to about 15 kV. An important candidate for momentum transfer across a closed magnetopause is the Kelvin-Helmholtz instability. Southwood (1979) suggested that, as the Kelvin-Helmholtz instability grows, nonlinear

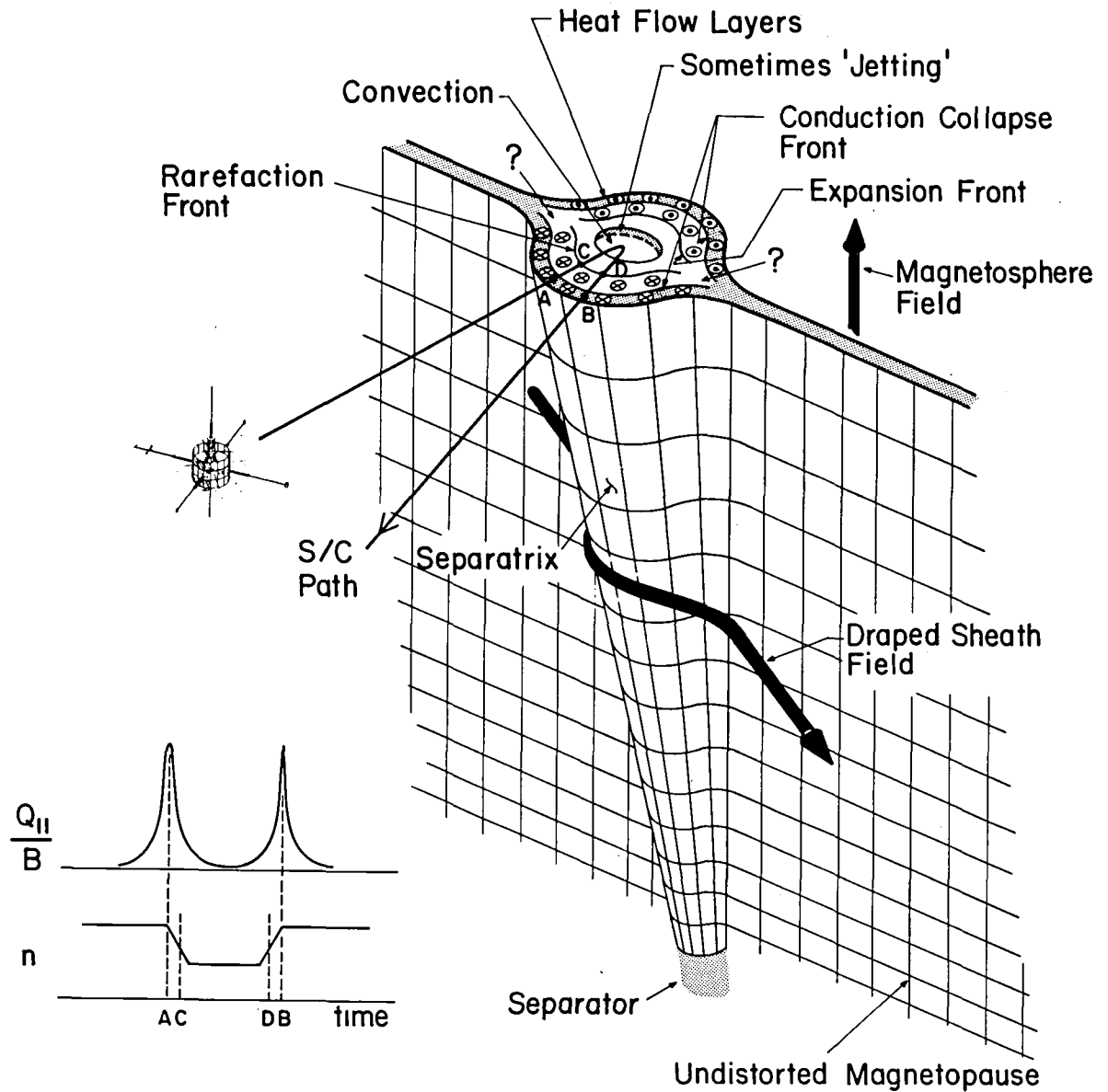


Figure 5-5. A schematic view of how flux transfer events might be interpreted as cases of solar wind magnetosphere coupling via magnetic field reconnection on the dayside magnetopause. (Adapted from Scudder et al., 1983.)

effects can add to net momentum and energy transport across the magnetopause. Pu and Kivelson (1983) have developed this idea quantitatively and describe the effects of this transport as well as the role of magnetopause surface waves in driving magnetic field-line resonances that can be detected on the ground as geomagnetic pulsations.

Once magnetosheath plasma gains entry into the magnetospheric boundary layer it then interacts with the geomagnetic field. In general, the boundary-layer plasma flow will have a finite cross-field component. By the ideal Ohm's law  $\underline{E} = -\underline{v} \times \underline{B}$ , a cross-field flow component will establish a polarization electric field independent of whether the boundary layer is on closed or open field lines. Figure 5-4 illustrates the electric fields and currents predicted in a simple open model (the Dungey model). This open model requires a load region with  $\underline{j} \cdot \underline{E} > 0$  near the dayside magnetopause (region 1) and near the tail current sheet (region 3). The tail magnetopause region (2) in turn acts as a generator region with  $\underline{j} \cdot \underline{E} < 0$  in which plasma energy is converted into electromagnetic energy. With a finite normal field component  $B_n$  near the tail magnetopause, this interaction can generate the dawn-dusk polar cap potential drop. As noted above the polar cap potential has a minimum value that may be associated with the  $B_n \rightarrow 0$  limit of a closed topology. In that case the boundary layer dynamo on closed field lines must generate the entire potential drop of  $\sim 15$  kV.

IMP, HEOS and ISEE satellite observations have shown that the magnetospheric boundary layer is supplied primarily by the magnetosheath although it also includes an ionospheric component. Magnetosheath plasma, momentum and energy generally decrease after crossing the magnetopause to produce the magnetospheric boundary layer. The boundary layer thus acts primarily as a generator region in which  $\underline{E} \cdot \underline{j} < 0$ , and hence plasma energy density contributes to the tail magnetic field energy and hence to Joule heating. If we consider region 2 of Figure 5-4 as being primarily a generator region, then essentially all of the outer magnetospheric boundary layer acts as a dynamo generator. The corresponding load regions are then the tail current sheet and the ionosphere.

Observations from the ISEE spacecraft of three-dimensional ion distributions in the low-latitude boundary layer on the sunward side (region 1) have shown that large cross-field flow components are commonly observed there. Direct sums over concurrently sampled electron and ion distributions also show the presence of field-aligned currents (Eastman et al., 1983). These results indicate that the magnetospheric boundary layer acts, at least in part, as a generator region for field-aligned currents that work to neutralize the polarization electric fields generated by the boundary layer plasmas. Direct electric field measurements from the ISEE spacecraft have just begun to provide information on the extent to which the outer magnetospheric boundary acts as a generator or load. More such *in situ* measurements of electric and magnetic fields and plasmas are needed to isolate the boundary conditions required for the global models discussed in Section IV.

The field-aligned currents generated in the boundary regions are illustrated in Figure 5-6 for both an open topology (5-6a) and a closed topology (5-6b). In either case plasmas near the outer magnetospheric boundary, whether on open or closed field lines, can act to produce currents of the same sense that link to the ionosphere. Thus, the current configuration away from the boundaries does not depend sensitively on exactly where the surface of last closed field lines is drawn.

As noted in Chapter 7 (Connections Between the Magnetosphere and Ionosphere), the existence of the tail current sheet, dividing the opposite hemispheres of the tail field, and the presence of a dawn-to-dusk cross-tail electric field indicate that significant energization will occur there.

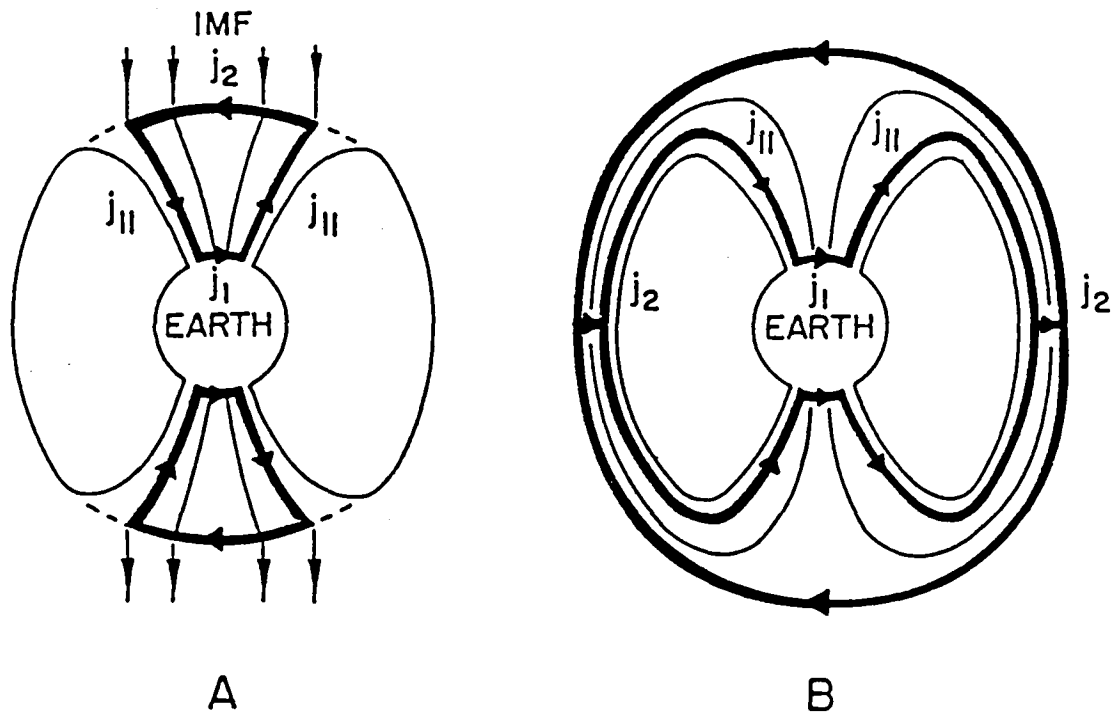


Figure 5-6. The "driving" current system (heavy lines) that transfers antisunward momentum to the polar cap ionosphere, directly from the solar wind in an open magnetosphere model (A) and from the low-latitude boundary layer in a closed magnetosphere (B). The view is from the Sun. (From Hill, 1983.)

The neutral sheet (region 3 in Figure 5-4) thus has  $\mathbf{j} \cdot \mathbf{E} > 0$ , and the associated energization has been estimated as  $3 \times 10^{11}$  to  $3 \times 10^{12}$  W (Stern, 1980). Charged particles may become trapped in the tail current sheet, execute oscillations about the sheet, and gain energy during the trapping interval from any quasi-steady electric field parallel (dawn-dusk direction) to the current sheet (e.g., Speiser, 1965). Lyons and Speiser (1982) numerically evaluated many particle orbits in a simple tail model and, using Liouville's theorem, mapped "initial" distributions through the acceleration process into "final" beam-like distributions on the edge of the model plasma sheet. These distributions compare quite well with observed distributions (e.g., DeCoster and Frank, 1979). In the parlance of solar wind-magnetosphere coupling, particles are fed in from the magnetospheric boundary layer and solar wind energy is fed in via the dawn-dusk electric field. The power dissipation discussed above is important. It is of the order of a few percent of the incident solar wind power on the magnetosphere.

Power dissipation due to steady-state convection at low altitudes and associated ionospheric heating is of a comparable magnitude (Hill and Wolf, 1977). For the generation region, the total available kinetic energy in the plasma mantle and low-latitude boundary layer is  $\approx 6 \times 10^{11}$  W (Eastman, 1979) which is again of the same order of magnitude as the estimate above.

## BOUNDARY CONDITIONS

Whatever entry mechanisms are thought to be dominant they need to be formulated in terms of boundary conditions for the plasma and the magnetic field at the magnetopause. This formulation must be achieved in order to connect the theories of solar wind entry with the theories of magnetospheric convection. Ideally, each entry mechanism should be reflected in a particular set

of boundary conditions to be specified over the entire surface of the magnetopause. In particular as a consequence of reconnection one must know the distribution of  $B_n(\underline{x})$  at the magnetopause. Moreover one must eventually be able to get transport coefficients provided either by kinetic theory or observations.

In order to illustrate the difficulties of formulating boundary conditions, let us consider a simplified version of the energy equation in the fluid picture.

$$-\frac{1}{8\pi} \frac{\partial}{\partial t} B^2 = \frac{1}{\sigma} j^2 + \frac{1}{4\pi} \nabla \cdot (\underline{E} \times \underline{B}) + \underline{v} \cdot (\underline{j} \times \underline{B}) \quad (2)$$

This equation shows that any time change in the magnetic field energy density at the magnetopause is associated with three different phenomena. The first term on the right-hand side describes dissipative Joule heating. This term is important if one considers a substantial violation of the frozen-flux condition (Equation 1) at the magnetopause. This is the case when solar wind entry mechanisms are considered in the context of a closed magnetosphere. The quantity  $1/\sigma$  is the resistivity. A model for the function  $1/\sigma$  must be derived from kinetic theory in order to take into account microphysical wave-particle interaction processes and/or microinstabilities operating near the magnetopause. These processes ultimately should lead to the formulation of a macroscopic resistivity function that will depend on position on the magnetopause surface.

The second term on the right-hand side is the divergence of the electromagnetic Poynting flux associated with plasma transport. The third term is the power density due to magnetic stresses parallel to the magnetosheath flow velocity vector. This term does not contribute to the change in the magnetic energy density when the magnetosphere is closed, i.e., when the magnetic normal component is zero at the boundary.

Let us assume for a moment a time-averaged steady-state situation in which the frozen-flux theorem (Equation 1) is justified. Then Equation (2) reduces to

$$\frac{1}{4\pi} \nabla \cdot (\underline{E} \times \underline{B}) = - \underline{j} \cdot \underline{E} \quad (3)$$

In this conventional form, the energy equation tells us that the conversion rate of magnetic to mechanical energy is equal to the divergence of the Poynting flux. To utilize even this simplified relation we need to know (a) the distribution of the Chapman-Ferraro currents on the magnetopause and (b) the distribution of the magnetic field component perpendicular to the magnetopause. The Chapman-Ferraro currents can be calculated by means of magnetostatics, provided the magnetic normal component is known. The magnetic normal component,  $B_n(\underline{x})$ , must be assumed a priori or derived from a macroscopic, time-averaged merging theory which we do not yet have.

First of all, one has to know where reconnection or merging occurs. Consequently, we must first develop a model for the location and shape of the "merging-line" or "separatrix" which separates regions of inward versus outward normal  $\underline{B}$ -field components. There are two prevailing hypotheses on the location of the merging line: (1) the "nose merging hypothesis" in which the merging line is assumed (sometimes tacitly) to pass through the subsolar point and to follow the direction of the magnetopause surface current; and (2) the "antiparallel merging hypothesis" in which the merging line is assumed to pass through the polar cusps and then to follow the locus of points at which the external and internal magnetic fields are exactly antiparallel. Neither hypothesis has yet been cast into a quantitative form that can be used in the prediction of the global distribution of the magnetic normal component.

#### IV. GLOBAL MODELING

The use of global MHD models for the magnetosphere is still in its early stages. The accuracy of the models so far is such that they serve to provide schematics that can be compared to theoretical ideas and to observations. There are a number of questions, including energy and momentum coupling to the magnetosphere from the solar wind, that demand a self-consistent global picture. The simplest such picture is the one provided by MHD. To the extent that the full particle distribution can be modeled by a few simple moments, the MHD picture is an accurate one. Many classes of kinetic effects can be included through a non-ideal formulation of the MHD equations with anomalous transport coefficients. Another aspect of the global models is the ability to monitor magnetospheric structure as a function of the varying solar wind input.

It is important to note that no global model has reached the point where boundary layers can be said to be calculated accurately, even assuming the adequacy of the MHD assumptions in these regions. In addition, the very assumptions of MHD can be shown to be violated in some boundary layers. However, it is the boundary layers that appear to control the physics of the magnetosphere and the solar wind-magnetosphere interaction. This lack is basically due to the problem of modeling the scale sizes of the boundary regions while simultaneously considering the global scale. This is a crucial problem for future work, but at present the boundary layers are poorly represented. This implies that the present generation of global models have built into them certain assumptions about coupling and entry processes. The first is that reconnection will occur whenever field directions are opposed. Brackbill (1982) has gone to some lengths to minimize this particular problem with some success. However, the largest effective magnetic Reynolds numbers,  $R_m$ , in these global models is still  $<100$  while the magnetosphere can have an  $R_m > 10^6$ . The second is that, due to numerical problems, viscous interactions at the boundary layers can be much stronger than physically warranted. Once again Reynolds numbers for these calculations tend to be  $<100$ . Finally, instabilities of the boundary layers (e.g., the Kelvin-Helmholtz instability at the magnetopause) are suppressed. These are important assumptions that are not always stated. They actually may be good assumptions that reflect the actual magnetosphere, but in any event, they do force the global models to produce certain types of results.

With these provisions the large-scale MHD simulations have aided considerably in the understanding of the physics of the magnetosphere. In discussing results of the global simulation efforts, the discussion will be split into two parts: coupling of the solar wind to the magnetosphere and magnetotail reconnection and convection.

There are three ways that the solar wind plasma can couple to the magnetosphere: pressure directed normal to a tangential discontinuity at the magnetopause, magnetic stresses produced at a rotational discontinuity, and viscous or quasi-viscous interactions.

The effect of pressure alone has been examined with the Chapman-Ferraro type of models. Wu (1983) has argued that the MHD nature of the solution in the interior of the magnetosphere fundamentally changes the structure of the magnetospheric field. Although the MHD model and Chapman-Ferraro model agree in the total magnetopause current distributions, the MHD model presents a very different physical picture from the Chapman-Ferraro model. In WU's MHD model with zero or northward IMF the cusp is not exposed directly to the solar wind because the day-side magnetosphere folds over the cusp region. Thus the solar wind plasma cannot enter the ionosphere directly through the cusp. Furthermore, in the MHD model a current sheet exists above the cusp region. The difference between the two models is attributed to the fact that the Chapman-Ferraro model is a vacuum model and thus only surface current is allowed. By having plasma in the dayside magnetosphere in the MHD model, a current sheet is allowed to form



inside the magnetopause just like the case for the tail current sheet. In any event, pressure forces on the magnetopause boundary alone do not produce any motion within the magnetosphere, at least in a steady-state situation. Non-steady forces of this type seem unlikely to produce anything resembling a simple two-cell convection pattern. Thus, it is left to magnetic stresses or viscous stresses to effectively couple the solar wind to the magnetosphere.

The magnetic stresses must ultimately be produced by reconnection between the IMF and the magnetospheric magnetic field. In those models where southward field components have been introduced in the solar wind, it appears that magnetic stresses are the dominant means of coupling. Brecht et al. (1981) have shown how eastward solar wind magnetic field components can be introduced into the magnetotail. In these simulations the magnetotail boundary layer seems to be relatively well represented in terms of density, velocity, thickness and temperature. This result really says nothing about the physics of the magnetopause; only that material can penetrate along flux tubes that cross the magnetopause boundary. In the cases where there is no merging of the solar wind magnetic field with the magnetosphere (i.e., no field or northward solar wind field), the effects of the viscous interaction are not yet sorted out from numerical effects. For example, the UCLA models (Wu et al., 1981; Wu, 1983) all seem to show fairly extended magnetotails similar to the observed magnetotail. In the NRL models, by contrast, the magnetotail contracts to a tear-drop shape and internal convection is much reduced.

The coupling of energy and plasma from the solar wind to the inner magnetosphere must be accomplished in the MHD picture through some variant of magnetospheric convection driven by dynamo processes acting near the magnetopause. This convection may either occur in a steady or non-steady fashion. In fact, there are good reasons for thinking that it must be more or less unsteady. It is in the maintenance of magnetospheric convection that tail reconnection is of primary importance. There have been a number of models for magnetotail reconnection, which might be fit into three categories: forced (Sato, 1979; Sato et al., 1983), spontaneous (Birn, 1980; Birn and Hones, 1981; Forbes and Priest, 1983), and global (Lyon et al., 1981; Brecht et al., 1982). A number of features are found in common with all of these models: plasma sheet thinning, jetting of the plasma at speeds on the order of the Alfvén speed, and plasma flow becoming more or less field-aligned. All of these features can easily be accommodated to the neutral line model of substorms. Further, all those models which make some attempt to include the proper magnetospheric boundary conditions show X-point formation in the near-Earth region of the tail and plasmoid formation. In addition, the 3-D models all tend to show that the neutral line is actually rather localized in cross-tail extent. This fact tends to bring the neutral line model in closer agreement with observations.

However, there are distinct problems with the reconnection models. The data do not show that all magnetospheric convection takes place with substorm-like events which are the only type of phenomenon that the reconnection simulations have as yet modeled. The reconnection models all have reconnection occurring close to the Earth. Attempts to force reconnection in the far magnetotail with a localized resistivity there lead to very unrealistic magnetosphere models which show field configurations in the near-Earth region that are highly unstable to tearing. Furthermore, some statistical correlations have been built up between substorm onset and antisunward flow (Hones and Schindler, 1979). The time scale of this observed tailward flow is relatively short. While some of the models show very sudden accelerations of plasma, they all show prolonged tailward flow. Lastly, the plasma sheet boundary layer shows up in none of these models. With MHD physics it is relatively easy to conceive of this boundary layer as being the signature of reconnection taking place further tailward. The persistence of the boundary layer for any length of time is a problem for MHD physics, since a normal  $\underline{B}$  component links it directly to the plasma sheet.

An important class of complementary global models are those that search for a steady-state equilibrium solution (e.g., Fuchs and Voigt, 1979). Here the structure of the solution is given by the specification of the magnetospheric boundaries and an internal pressure function. Since in these models the boundaries are inputs, there are essentially no numerical artifacts associated with the boundary layer. They then provide an important check on the structure shown by the time-dependent models. Even time-dependent behavior can be modeled as a sequence of equilibria for time scales longer than those associated with substorms.

While global models have so far been quite successful as conceptual aids, a great deal of work needs to be done before they can be guides rather than aids. Of primary importance here is the treatment of boundary layers where the coupling between various regions occurs. Almost any type of physics can in some way be incorporated into an MHD model. What is necessary is that the physics be elucidated so that it can be included in the simulation models.

## V. SUMMARY AND OUTSTANDING PROBLEMS

A simple assessment of our present understanding of solar wind-magnetosphere coupling can be made in terms of the four major questions that were posed in Section I. The first of these questions is concerned with the process of solar wind plasma entry. Although solar wind plasma entry is not totally understood we do have observational evidence for the efficient injection of plasma through the cusps and into the plasma mantle. The open magnetic topology which is likely associated with reconnection and flux transfer events is at least partially responsible for cusp entry and possibly also for some plasma transport into the low-latitude boundary layer. Impulsive plasma injection also appears to be a quite reasonable hypothesis for the localized efficient injection of plasma into the low-latitude boundary layer. Finally, a number of cross-field diffusion processes have been proposed as closed model entry processes.

Further progress on plasma entry will require energetic particle and plasma measurements with high time resolution and high resolution in phase space, coupled with electric and magnetic field and plasma wave measurements with good coverage of the dayside magnetopause. Full 3-D plasma and particle distribution measurements are especially needed in the region of the outer cusp and the dayside boundary layer poleward of the cusp. Moreover, continuous measurements of the properties of the upstream solar wind in the vicinity of the Earth are critically important for future experimental studies of the solar wind-magnetosphere interaction. The different models that have been proposed to explain plasma entry must be refined into theories that can be tested by plasma and field measurements. Since reconnection is a major candidate process, suitable modifications to the original Petschek (1964) theory should be made to treat the three-dimensional aspects of the solar wind interactions and to examine the kinetic theory of the reconnection diffusion regions.

The second question concerns the processes of energy and momentum transfer from the solar wind to the magnetosphere. In this area we know that both open- and closed-model processes are important. The open models that result from steady-state reconnection and flux transfer events readily provide qualitative explanations for the response of magnetospheric convection and substorm activity to changes in solar wind parameters. On the other hand, closed-model processes are needed to explain the minimum polar cap potential drop of some 15 kV. This minimum potential is likely associated with the low-latitude boundary layer acting as an MHD generator on closed magnetic field lines. The suitability of the impulsive plasma injection model should be more fully explored and tested against spacecraft plasma measurements near the magnetopause. ISEE satellite measurements of waves and structure near the magnetopause provide valuable new information relevant to momentum transport via mechanisms such as the Kelvin-Helmoltz instability.

The third question deals with the physics of magnetospheric boundary layers, specifically their role as generators, loads and plasma transport regions. A major result of recent spacecraft programs has been the establishment of the existence and importance of these boundary layers. A more advanced level of understanding of magnetospheric boundary layers requires high-resolution measurements of their structure and dynamics. Theories of boundary-layer formation need to account for both MHD and kinetic effects and to establish the relationship of these effects to time-varying solar wind parameters.

The final question concerns the global magnetohydrodynamics that characterize the magnetosphere for the various coupling processes and as functions of solar wind parameters. Already the MHD models are capable of approximately describing the coupling between various magnetospheric regimes. As a complement to the global MHD models, magnetostatic equilibrium models are capable of describing the configuration of the quiet magnetosphere with higher spatial resolution. Both approaches to global modeling have now been extended to treat the system in three dimensions. Development of complete models will require an enhanced ability to specify coupling processes and boundary conditions such as the plasma distributions and current systems that characterize the various magnetospheric boundary layers.

The variety of magnetospheric phenomena discussed in this chapter leads to a number of unresolved physical questions that need to be answered in the light of observations, as well as from a theoretical point of view. In order to advance our present understanding, we recommend consideration of the following items:

1. There is observational evidence that the plasma near the magnetopause boundary, and near the plasma sheet boundary as well, is often highly anisotropic. Moreover, we know that some plasma drift mechanisms ultimately lead to an energy dispersion of the thermal plasma. Consequently, global MHD models for the magnetosphere, and magnetohydrostatic equilibrium models as well, should be formulated eventually in terms of multifluid codes and in terms of anisotropic thermal plasma pressure.

2. There is an overall agreement that the interaction between the plasma and the magnetic field within discontinuity regions (e.g., the magnetopause and the plasma sheet boundary layer) is dominated by microphysical plasma processes. The global fluid models need macroscopic physical quantities in their basic equations. It is, therefore, necessary to have a global model for the resistivity at the magnetopause. Such a model could be obtained from observations or derived from the kinetic theory. For example, information is essential for a proper calculation of the magnetic normal component  $B_n(\underline{x})$  at the magnetopause.

3. At the magnetopause, the quantity  $\underline{j}(\underline{x}) \cdot \underline{E}(\underline{x})$  is a measure of the conversion rate of magnetic into mechanical energy and vice versa. The Chapman-Ferraro current density  $\underline{j}(\underline{x})$  is determined uniquely by the distribution of the magnetic normal component  $B_n(\underline{x})$  at this boundary. Is it possible to derive macroscopic boundary values  $B_n(\underline{x})$  on the basis of kinetic arguments (see item 2) in order to calculate the global energy conversion rate at the magnetopause?

4. The magnetic normal component  $B_n(\underline{x})$  determines the "openness" of the magnetosphere, the size of the auroral oval, and the ionospheric convection  $\underline{E}$ -field pattern. The polar cap electric field is assumed to reflect the solar wind  $(\underline{v} \times \underline{B})$  field under the assumption that the ideal MHD theory is applicable. We know, however, that there exist significant magnetic field-aligned electric fields along the auroral oval. To what extent do field-aligned electric fields affect the polar cap convection  $\underline{E}$ -field which is mapped from the solar wind through the magnetopause down to the ionosphere?

5. Sources for the plasma sheet are thought to include the ionosphere, the low-latitude boundary layer, and the plasma mantle. Are these known plasma sources sufficient to account for plasma sheet replenishment on substorm time scales?

6. Can we improve measurements and understanding of the minimum polar cap potential during extended quiet periods with northward IMF?

7. From the standpoint of global modeling as well as from our understanding of the time-dependent reconnection process, it is vital to obtain improved kinetic theory of the diffusion regions at the reconnection site.

8. Are flux transfer events the "debris" or active signatures of reconnection? Are they possibly associated with "impulsive injection events" or finite plasma filaments?

9. Can statistical analyses using solar wind and magnetospheric parameters be used to compare the relative efficacy of driven versus indirectly driven models of magnetospheric response and of closed versus open models of coupling? The results of such studies could act as a guide to the development of better theoretical models of solar wind-magnetospheric coupling.

10. Does the average plasma flow in the plasma sheet correspond to the predictions of MHD convection models? Observations of this flow and of  $B_z$  can give the pattern of the convection electric field. This pattern is likely to be quite different for open- and closed-coupling processes. Thus, these observations can provide a check on the importance of both processes as well as on the validity of convection models.

11. A tangential component of the electric field at the magnetopause is predicted as a consequence of magnetic field reconnection, and thus electric field measurements at the magnetopause are especially important. Special efforts should be made to develop and apply improved definitive measurements of vector electric fields at the magnetopause.

In Figure 5-1 we have illustrated the various plasma regimes of the Earth's magnetosphere. Our evaluation of the solar wind-magnetosphere coupling problem leads us to emphasize the importance of the boundary layers as the primary transport and coupling regions for the magnetosphere even though they constitute <5% by volume of the total system. We recommend that future observational and theoretical studies should focus on these boundary regions to determine the boundary conditions that they impose on the magnetospheric system and to isolate the dominant coupling processes.

## VI. REFERENCES

- Aggson, T. L., P. J. Gambardella, and N. C. Maynard, Electric field measurements at the magnetopause: 1. Observations of large convective velocities at rotational magnetopause discontinuities, submitted to J. Geophys. Res., 1983.
- Axford, W. I., and C. O. Hines, A unifying theory of high latitude geophysical phenomena and geomagnetic storms, Can. J. Phys., **39**, 1433, 1961.

- Baker, D. N., R. D. Zwickl, S. J. Bame, E. W. Hones, B. T. Tsurutani, E. J. Smith, and S. -I. Akasofu, An ISEE-3 high time resolution study of interplanetary parameter correlations with magnetospheric activity, J. Geophys. Res., in press, 1983.
- Bargatze, L. F., D. N. Baker, R. L. McPherron, and E. W. Hones, Magnetospheric response for many levels of geomagnetic activity, J. Geophys. Res., submitted, 1983.
- Birn, J., Computer studies of the dynamic evolution of the geomagnetic tail, J. Geophys. Res., **85**, 1214, 1980.
- Birn, J., and E. W. Hones, Three-dimensional computer modeling of dynamic reconnection in the geomagnetic tail, J. Geophys. Res., **86**, 6802, 1981.
- Brecht, S. H., J. G. Lyon, J. A. Fedder, and K. Hain, A simulation study of east-west IMF effects on the magnetosphere, Geophys. Res. Lett., **8**, 397, 1981.
- Brecht, S. H., J. G. Lyon, J. A. Fedder, and K. Hain, A time dependent three-dimensional simulation of the earth's magnetosphere: reconnection events, J. Geophys. Res., **87**, 6098, 1982.
- Burch, J. L., P. H. Reiff, and M. Sugiura, Upward electron beam measured by DE-1: a primary source of dayside region-1 Birkeland currents, Geophys. Res. Lett., **10**, 753, 1983.
- Brackbill, J. U., Numerical modeling of magnetospheric reconnection, Los Alamos National Lab. preprint LA-UR-82-483, 1982.
- DeCoster, R. J., and L. A. Frank, Observations pertaining to the dynamics of the plasma sheet, J. Geophys. Res., **84**, 5099, 1979.
- Dungey, J. W., Interplanetary magnetic field and the auroral zones, Phys. Rev. Lett., **6**, 47, 1961.
- Eastman, T. E., The plasma boundary layer and magnetopause layer of the earth's magnetosphere, Ph.D. thesis LA-7842-T, Los Alamos Scientific Lab., 1979.
- Eastman, T. E., L. A. Frank, and C. Y. Huang, The boundary layers as the primary transport regions of the earth's magnetotail, submitted to J. Geophys. Res., 1983.
- Eviatar A., and R. A. Wolf, Transfer processes in the magnetopause, J. Geophys. Res., **73**, 5561, 1968.
- Fuchs, K., and G. -H. Voigt, Self consistent theory of a magnetospheric B-field model, in Quantitative Modeling of Magnetospheric Processes, ed. by W. P. Olson, p. 86, AGU, Washington, D.C., 1979.
- Forbes, T. G., and E. R. Priest, A numerical experiment relevant to line-tied reconnection in two-ribbon flares, Solar Phys., **84**, 169, 1983.
- Gary, S. P., and T. E. Eastman, The lower hybrid drift instability at the magnetopause, J. Geophys. Res., **84**, 7378, 1979.
- Gary, S. P., and S. J. Schwarz, An electrostatic flow shear instability, J. Geophys. Res., **85**, 2978, 1980.

- Gosling, J. T., J. R. Asbridge, S. J. Bame, W. C. Feldman, G. Paschmann, N. Sckopke, and C. T. Russell, Evidence for quasistationary reconnection at the dayside magnetopause, J. Geophys. Res., **87**, 2147, 1982.
- Hasegawa, A., and K. Mima, Anomalous transport produced by kinetic Alfvén wave turbulence, J. Geophys. Res., **83**, 1117, 1978.
- Haerendel, G., Microscopic plasma processes related to reconnection, J. Atmos. Terr. Phys., **40**, 343, 1978.
- Hill, T. W., Solar-wind magnetosphere coupling in Solar Terrestrial Physics: Principles and Theoretical Foundations, ed. by R. Carovillano and J. M. Forbes, D. Reidel (Boston) in press, 1983.
- Hill, T. W., and R. A. Wolf, Solar-wind interactions, in The Upper Atmosphere and Magnetosphere, National Academy of Sciences, Washington, D.C. 1977.
- Hones, E. W., and K. Schindler, Magnetotail plasma flow during substorms: a survey with Imp 6 and Imp 8 satellites, J. Geophys. Res., **84**, 7155, 1979.
- Lemaire, J., Impulsive penetration of filamentary plasma elements into the magnetospheres of the earth and Jupiter, Planet. Space Sci., **25**, 887, 1977.
- Lemaire, J., and M. Roth, Penetration of solar wind plasma elements into the magnetosphere, J. Atm. Terr. Phys., **40**, 331, 1978.
- Levy, R. H., H. E. Petschek, and G. L. Siscoe, Aerodynamic aspects of the magnetospheric flow, AIAA J., **2**, 1965, 1964.
- Lyon, J. G., S. H. Brecht, J. D. Huba, J. A. Fedder, and P. J. Palmadesso, Computer simulation of a geomagnetic substorm, Phys. Rev. Lett., **46**, 1038, 1981.
- Lyons, L. R., and T. W. Speiser, Evidence for current sheet acceleration in the geomagnetic tail, J. Geophys. Res., **87**, 2276, 1982.
- Paschmann, G., B. U. Ö. Sonnerup, I. Papamastorakis, N. Sckopke, G. Haerendel, S. J. Bame, J. R. Asbridge, J. T. Gosling, C. T. Russell, and R. C. Elphic, Plasma acceleration at the earth's magnetopause: evidence for reconnection, Nature, **282**, 243, 1979.
- Paschmann, G., G. Haerendel, I. Papamastorakis, N. Sckopke, S. J. Bame, J. T. Gosling, and C. T. Russell, J. Geophys. Res., **87**, 2159, 1982.
- Perreault, P., and S. -I. Akasofu, A study of geomagnetic storms, Geophys. J. Roy. Ast. Soc., **54**, 547, 1978.
- Petschek, H. E., Magnetic field annihilation, in The Physics of Solar Flares, ed. by W. N. Hess, p. 425, NASA, Washington, D.C., 1964.
- Pu, Z. Y., and M. G. Kivelson, Kelvin-Helmholtz instability at the magnetopause: energy flux into the magnetosphere, J. Geophys. Res., **88**, 853, 1983.
- Reiff, P. H., R. W. Spiro, and T. W. Hill, Dependence of polar cap potential drop on interplanetary parameters, J. Geophys. Res., **86**, 7639, 1981.

- Russell, C. T., and R. C. Elphic, Initial ISEE magnetometer results: magnetopause observations, Space Sci. Rev., **22**, 681, 1978.
- Russell, C. T., and R. C. Elphic, ISEE observations of flux transfer events at the dayside magnetopause, Geophys. Res. Lett., **33**, 1979.
- Sato, T., Strong plasma acceleration by slow shocks resulting from magnetic reconnection, J. Geophys. Res., **84**, 7177, 1979.
- Sato, T., T. Hayashi, R. J. Walker, and M. Ashour-Abdalla, Neutral sheet current interruption and field-aligned current generation by three-dimensional driven reconnection, Geophys. Res. Lett., **10**, 221, 1983.
- Schindler, K., On the role of irregularities in plasma entry into the magnetosphere, J. Geophys. Res., **84**, 7257, 1979.
- Scudder, J. D., K. W. Ogilvie, and C. T. Russell, The electron plasma signatures of flux transfer intervals: Evidence for ongoing, localized, time-dependent reconnection, submitted to J. Geophys. Res., 1983.
- Sonnerup, B. U. Ö., G. Paschmann, I. Papamastorakis, N. Sckopke, G. Haerendel, S. J. Bame, J. R. Asbridge, J. T. Gosling, and C. T. Russell, Evidence for magnetic field reconnection at the earth's magnetopause, J. Geophys. Res., **86**, 10049, 1981.
- Southwood, D. J., Magnetopause Kelvin-Helmholtz instability, in Magnetospheric Boundary Layers, ed. by B. Battrock, European Space Agency special publ. ESA SP-148, 1979.
- Speiser, T. W., Particle trajectories in model current sheets, 1. Analytical solutions, J. Geophys. Res., **70**, 4219, 1965.
- Speiser, T. W., and D. J. Williams, Magnetopause modeling: flux transfer events and magnetosheath quasi-trapped distributions, J. Geophys. Res., **87**, 2177, 1982.
- Speiser, T. W., D. J. Williams, and H. Garcia, Magnetically trapped ions as a source of magnetosheath energetic ions, J. Geophys. Res., **86**, 723, 1981.
- Stern, D. P., Energetics of the magnetosphere, NASA/GSFC TM 82039, 1980.
- Tsurutani, B. T., and R. M. Thorne, Diffusion processes on the magnetopause boundary layer, Geophys. Res. Lett., **9**, 1247, 1982.
- Vasyliunas, V. M., Theoretical models of magnetic field line merging I., Rev. Geophys. Space Phys., **13**, 303, 1975.
- Wu, C. -C., Shape of the magnetosphere, submitted to Geophys. Res. Lett., 1983.
- Wu, C. -C., R. J. Walker, and J. M. Dawson, A three dimensional MHD model of the earth's magnetosphere, Geophys. Res. Lett., **8**, 523, 1981.
- Wygant, J. R., R. B. Torbert, and F. S. Mozer, Comparison of S3-3 polar cap potential drop with the interplanetary magnetic field and models of magnetopause reconnection, J. Geophys. Res., **88**, 5727, 1983.





## CHAPTER 6

# CORONAL TRANSIENTS AND THEIR INTERPLANETARY EFFECTS

### WORKING GROUP MEMBERS

A. J. Hundhausen, Chairman  
*National Center for Atmospheric Research*

L. F. Burlaga  
*NASA/Goddard Space Flight Center*

W. C. Feldman  
*Los Alamos National Laboratory*

J. T. Gosling  
*Los Alamos National Laboratory*

E. Hildner  
*NASA/Marshall Space Flight Center*

L. L. House  
*National Center for Atmospheric Research*

R. A. Howard  
*Naval Research Laboratory*

A. S. Krieger  
*American Science and Engineering, Inc.*

M. R. Kundu  
*University of Maryland*

B. C. Low  
*National Center for Atmospheric Research*

N. R. Sheeley, Jr.  
*Naval Research Laboratory*

R. S. Steinolfson  
*University of California, Irvine*

R. T. Stewart  
*Commonwealth Scientific and Industrial Research Organization*

R. G. Stone  
*NASA/Goddard Space Flight Center*

S. T. Wu  
*University of Alabama*

**CHAPTER 6**  
**CORONAL TRANSIENTS AND THEIR INTERPLANETARY EFFECTS**

I.	Introduction	6-3
II.	Background Material	6-4
III.	The Present	6-10
IV.	The Future	6-20
V.	References	6-24

## I. INTRODUCTION

A statistical association between solar flares and geomagnetic storms has been widely accepted since early in the 20th century, and the interpretation of this effect in terms of an expulsion of material from the Sun has been a commonplace in the field of solar-terrestrial physics for nearly as long. In situ observations of outward-propagating interplanetary shock waves made in the 1960s provided direct evidence for such a phenomenon. The masses inferred from the shock wave observations implied that their passage through the solar corona should produce major perturbations of that region. Thus the first identifications of coronal transients or mass ejections by spacecraft-borne instruments in the early 1970s were of interest to a far broader group of scientists than those who specialized in this most tenuous domain of solar physics. The role of coronal mass ejections in solar-terrestrial physics, or more specifically, in the chain of cause and effect introduced above (solar flare-interplanetary shock wave-geomagnetic storm), has remained an important, continuing theme in the study of mass ejections. Analysis of the "first generation" of these coronal observations, development of theoretical models of mass ejections, and comparisons with both solar and interplanetary data have yet to provide definitive answers to the first two questions that arise in pursuit of this theme: (1) What are the solar origins (in both a phenomenological and physical sense) of coronal mass ejections, and (2) What are the interplanetary effects of coronal mass ejections?

These two fundamental and familiar questions concerning the role of coronal mass ejections in the context of solar-terrestrial physics were the foci of our discussions at this workshop. In this document, we first describe some of the necessary background material on flares and geomagnetic storms, on interplanetary shock waves, and on coronal mass ejections themselves. We then describe one of the modern "tools" available for approaching these questions -- theoretical models for the initiation and propagation of transient phenomenon in the solar corona. All of this material has been extensively reviewed in the recent literature, and our coverage of it will be both selective and somewhat abbreviated. We then describe the second of our "tools" -- a new generation of coronagraph observations of mass ejections and complementary set of solar and interplanetary observations suitable for correlative studies. Much of this material has been published only recently, is available in preprint form, or resides on the desks and in the notebooks of those attending this workshop.

We will then conclude with a discussion of important problems to be solved in answering the two fundamental questions posed above and of some suggested strategies for approaching these problems. Before launching into this discussion, it may be appropriate to repeat a warning concerning work on this very topic stated by one of the pioneers of solar-terrestrial physics, C. Chree, in 1912.

*There is no side of Terrestrial Magnetism which appeals so much to the imagination as magnetic storms, and none presents equal fascination for theorists. It is difficult to get the ordinary man - even the ordinary scientific man - to recognize that in this, as in other cases, minute study of the facts ought desirably to precede speculation as to the cause of the phenomena.*

## II. BACKGROUND MATERIAL

### ANCIENT HISTORY—SOLAR FLARES AND GEOMAGNETIC STORMS

On September 1, 1859, the British astronomer Carrington, while engaged in his "...customary observation of the forms and positions of the solar spots,..." witnessed a short-lived brightening of a small part of the solar photosphere. A second diligent recorder of sunspots, Hodgson, noted the same event independently from a different observing site. This was the first recorded detection of the rare phenomenon now known as a white-light solar flare. In their simultaneous announcements to the Royal Astronomical Society (Carrington, 1860; Hodgson, 1860), both observers noted that a moderate disturbance of the geomagnetic field occurred at virtually the same time as their unusual solar event and that a major magnetic storm commenced 16 hours later. Despite statement of proper qualifications (in Carrington's words, "One swallow does not make a summer"), the implication of a solar cause and terrestrial effect was clear. This possible relationship between solar flares and geomagnetic storms became one of the three foundations of early solar-terrestrial physics (the others: 27-day recurrence tendency and 11-year variations in geomagnetic activity).

A thorough review of the historical development of this idea would not be appropriate here. Suffice it to say that by the mid-20th century the general solar origin of geomagnetic activity was well established, the interplanetary "transmitter" of the major effects was recognized to be particles rather than electromagnetic radiation, and the flare-storm association was attributed to the explosive ejection of a cloud of ionized solar material by the flare. Two crucial steps in the development of solar-terrestrial physics might well serve as object lessons in our forthcoming discussion. Much confusion in establishing the specific solar phenomena responsible for geomagnetic activity remained until it was recognized that there were two different classes of that activity, large, sporadic storms and small or moderate recurrent disturbances (e.g., Greaves and Newton, 1929) and that these classes had different solar origins. Much effort was expended (see Chapters 5, 11, and 12 of Chapman and Bartels, 1940) in following the relationship between sunspots and geomagnetic activity, suggested by the 11-year variation in both, to a dead end. The lack of any physical association between individual sunspots and geomagnetic disturbances seems to have been demonstrated many times but accepted slowly and reluctantly.

The interpretation of the flare-storm association in terms of an ejected cloud of material required only minor modification when it was recognized (in the late 1950s) that interplanetary space was permeated by a continuous flow of plasma from the Sun, the solar wind; the flare need only produce a short-lived burst of fast or energetic solar wind whose arrival at the Earth would perturb the "quiet" state of the magnetosphere. In fact, this modification led to a simple explanation of the sudden commencement of most flare-associated geomagnetic storms. In contrast to a cloud moving through a near vacuum, a perturbation of a fluid-like medium could have a sharp leading edge if it were sufficiently strong to form a shock wave. This suggestion was so intriguing that it prevailed despite the objection that any reasonable estimate of the collision rate in the solar wind near the orbit of Earth was so low that an ordinary shock front should not form; the "collision-free" shock (Gold, 1955) was originated to overcome this difficulty.

This concept was developed and to some extent refined before in situ solar wind observations became available. It led both to the theoretical study of shock propagation through the solar wind (Parker, 1961, 1963) and quantitative empirical studies of the still hypothetical ejection of material by solar flares. Results from both of these categories proved to be of great value to the interpreters of early solar wind observations and led to quantitative understanding that has remained valid to this day. For example, Akasofu and Yoshida (1967) found a typical delay time

between associated flares and geomagnetic storm sudden commencements of  $\sim 40$  hours. If interpreted as a Sun-Earth transit time, this delay implied an average transit speed for the cloud or shock of  $\sim 1000$  km sec $^{-1}$ . The same authors showed that flares within about  $60^\circ$  in longitude of the central meridian of the solar disc seemed to produce geomagnetic effects. That is, the ejected cloud and shock wave moved outward with a broad front rather than as a narrow jet.

## MODERN HISTORY — INTERPLANETARY SHOCK WAVES

In the early 1960s plasma probes and magnetometers carried beyond the terrestrial magnetosphere confirmed the existence of collision-free shocks moving outward from the Sun. These shocks occurred at the leading edge of "waves" of abnormally fast solar wind, the fluid conditions that would be expected to have steepened to form a shock front. Although these shock waves produced major enhancements of mass, momentum, and energy fluxes carried by the solar wind, the shocks occurred so infrequently and the enhancements were of sufficiently short duration that their influence on average solar wind fluxes were small. Well behind the front, abrupt changes in the chemical composition and magnetic field of the solar wind were observed and interpreted as the arrival of material with a different origin than the "swept-up" solar wind in the compressed shell; this material was regarded as the coronal or chromosphere plasma originally ejected by the flare, or the "driver gas" for the entire solar wind disturbance.

The geomagnetic effects of the observed interplanetary shock waves were much as anticipated from the solar-terrestrial studies summarized above. The sudden jump in momentum flux associated with shock passage was found to compress the magnetosphere, producing the sudden commencement and the initial phase of a geomagnetic storm. However, the relationship of the observed interplanetary shock waves to solar flares was less clear than expected. While many interplanetary shock waves were obviously associated with major flares, as then identified and classified by H $\alpha$  flare patrols, irritating exceptions were all too common. Some interplanetary shock waves had no reasonable flare associations and many major solar flares produced no discernible interplanetary disturbances despite being favorably located near the center of the visible solar disc.

By the early 1970s a sufficient number of interplanetary shock waves had been observed and studied to justify a synthesis of their characteristics into a description of the "typical" shock disturbance in the solar wind. Figure 6-1 shows one attempt at such a synthesis (Hundhausen, 1972). It includes some features of the "post-shock" plasma flow and chemistry mentioned above. It once again shows a "broad front" geometry for the disturbance similar to that inferred by studies of flare-storm associations; however this shock shape could be based upon an additional piece of evidence. Measurement of plasma and magnetic properties on both sides of a shock can be used to determine the local shock speed and the direction of the normal to the local shock surface through the Rankine-Hugoniot relations (e.g., Colburn and Sonett, 1966). Similar observations made by two closely-spaced spacecraft allowed still more accurate determinations of these shock properties. Observed shock normals clustered about the direction radial from the Sun with a spread of  $20\text{--}30^\circ$ . The shock fronts were not Sun-centered spheres (all normals would then be radial), but were broad. Taylor (1969) and Hirshberg (1968) used both observed shock normals and flare-sudden commencement associations to infer that the radius of curvature of the typical shock front near the orbit of Earth was about  $3/4$  of an astronomical unit. The observed structure of the post-shock solar wind flow was usually consistent with a shock driven by a lengthy emission of fast solar wind that reaches the orbit of Earth along a broad front rather than a narrow jet.

The local speeds of shock waves observed near the orbit of Earth were found to be typically  $500\text{--}600$  km sec $^{-1}$ ; speeds as high as  $1000$  km sec $^{-1}$  have been found only rarely. Since the

POSSIBLE GEOMETRY OF A FLARE - PRODUCED SHOCK WAVE,  
SYNTHESIZED FROM OBSERVATIONS

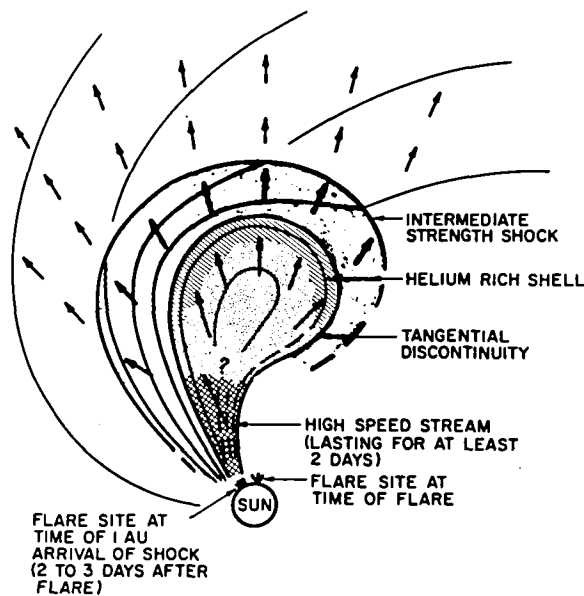


Figure 6-1. A Synthesis of interplanetary shock observations and their interpretations, circa 1970 (Hundhausen, 1972).

ambient solar wind has an average speed near  $400 \text{ km sec}^{-1}$ , the typical interplanetary shock was moving at only  $100\text{--}200 \text{ km sec}^{-1}$  with respect to the ambient; the shocks are of intermediate strength and are as much "convected" along in the overall flow of the solar wind as propagating through it. Comparison of the inferred local shock speeds with transit speeds deduced from flare associations generally implied a mild deceleration of the shock in transit from the Sun to the orbit of Earth,  $V_{\text{local}}/V_{\text{transit}} \approx 1/2$  to 1. More recent measurements by spacecraft well within or beyond the orbit of Earth confirm that most interplanetary shock waves move with nearly constant or slowly declining speeds.

The features of Figure 6-1 that may be most important to our discussion of coronal mass ejections are the content and magnetic topology of the broad cloud of material ejected by the flare. The question mark near the sunward end of the stipled region on Figure 6-1 indicated the uncertain connection of the magnetic field within the ejecta to its source at the Sun. This question arose very early in the theoretical discussion of plane-associated interplanetary shock waves. In his quantitative models of shock wave propagation, Parker (1961) emphasized the compression of spiral magnetic field lines in the ambient solar wind by the shock wave and drew a similar spiral field structure within the fast-moving, post shock solar wind. In less quantitative descriptions of the magnetic fields resulting from flares, Gold (1966) suggested that magnetic tongues or bottles were blown outward into interplanetary space. This latter configuration then suggested that the distended magnetic loops in the bottle might be "cut-off", as in Figure 6-2, to form closed structures (plasmoids) that were disconnected from the Sun. At the time Figure 6-1 was drawn there was no convincing observational evidence to distinguish among the possibilities that the magnetic field lines in the flare ejecta were simply distorted spiral field lines, distended loops still rooted in Sun, or closed, detached structures.

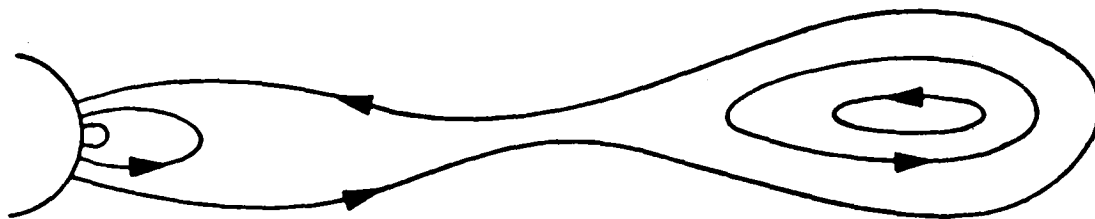


Figure 6-2. The hypothetical "cut-off" of an arcade of distended magnetic loops, as might be formed in the ejection of material by a solar flare, to form a detached magnetic structure in the solar wind (Gold, 1966).

Integration of the solar wind mass and energy flux densities over the time during which an interplanetary shock wave sweeps past an observer yields estimates of the mass and energy per unit area added to the solar wind by this disturbance. Adoption of a shock wave geometry similar to that shown in Figure 6-1 then leads to an estimate of the total mass and energy (above that in the ambient flow) carried by the shock wave. Estimated masses have ranged from  $5 \times 10^{15} \text{ g}$  to  $1.5 \times 10^{17} \text{ g}$ , and estimated energies have ranged from  $5 \times 10^{30} \text{ erg}$  to  $5 \times 10^{32} \text{ erg}$ . A typical shock wave mass is  $\sim 3 \times 10^{16} \text{ g}$ , a typical energy is  $\sim 10^{32} \text{ erg}$ . These values imply that the perturbation of the corona produced by the passage of a flare-associated shock wave must be large. For example, the normal solar wind carries  $3 \times 10^{16} \text{ g}$  out of the entire corona on a time scale of about eight hours; a flare produced shock wave carries a similar mass out of a limited region of the corona on a shorter time scale.

Thus the body of observations of interplanetary shock waves accumulated by the early 1970s was largely consistent with the expectations based on earlier studies of solar-terrestrial physics ("Ancient History" of this section). Some solar flares clearly did eject material that traveled through interplanetary space, formed a shock wave, and produced a magnetic storm through interaction with the terrestrial magnetosphere. Although arguments such as that given above suggested large effects in the solar corona, coronal observations available at that time were made with such poor temporal resolution that flare or shock-associated modifications of the coronal brightness (or density) structure were not directly observed while moving through the corona. The main evidence for such phenomena in the corona came, rather, from observations of metric radio emission. Two of the classical forms (see the reviews by Wild et al., 1963 and Wild and Smerd, 1972) of emission were interpreted as evidence for outward shock propagation (type II radio bursts) and for outward motion of a plasmoid (moving type IV emission). The suggested

relationship to the flare interplanetary shock wave phenomenon was strengthened by the association of type II and type IV radio emission with flares, interplanetary shock waves, and geomagnetic activity (e.g., Kundu, 1965). The first interplanetary type II radio burst (at kilometric wavelengths), observed by Malitson et al. (1973), traced a shock wave from a heliocentric distance of  $\sim 50$  solar radii to the orbit of the Earth at an average speed greater than  $1200 \text{ km sec}^{-1}$ . The shock speeds derived from these radio observations using a reasonable coronal density model were usually  $1000\text{--}1500 \text{ km sec}^{-1}$  (e.g., Wild et al., 1963).

## CORONAL TRANSIENTS OR MASS EJECTIONS

The first direct observations of a coronal phenomenon suggesting the expected passage of material ejected from a flare were made by spacecraft-borne coronagraphs in the early 1970s. Such instruments flown on OSO-7 (Tousey, 1973; Howard et al., 1976) and Skylab (MacQueen et al., 1974) detected outward moving, bright coronal features. These discrete expulsions of coronal material have often been called "coronal transients;" we will use the more restrictive term "coronal mass ejections" (Gosling et al., 1974a; Munro et al., 1979; Hundhausen et al., 1983) to distinguish events that clearly show ejection of coronal mass from the myriad other time-dependent phenomena (e.g., motion or brightening of coronal rays or streamers) observable with these and more recent coronagraphs. Reviews of the early observations of mass ejections and of their relationship to other solar and interplanetary phenomena have been given by Gosling (1975), Hildner (1977), MacQueen (1980), Rust et al., (1980), and Rust (1983). The related radio observations have been reviewed by Dulk (1980) and Stewart (1980). Thus a brief recapitulation of the characteristics of mass ejections, with emphasis on their role as the possible coronal manifestations of flare-associated shock waves, is appropriate here.

The mass ejections most often described (see, however, the section entitled "Association Between Coronal Mass Ejections and Interplanetary Phenomena" in Section III) in the literature are large, bright features seen in successive coronagraph images to rise above the limb of the Sun. These bright ejections span ranges of position angle (measured around the Sun from the north rotational pole) as large as  $120^\circ$ ; a typical angular spread is about  $50^\circ$ . Many mass ejections have the appearance of bright "loops"; about  $1/3$  of the mass ejections seen by the Skylab coronagraph were classified as loop-like (Munro et al., 1979). It must be emphasized, however, that this description carries no implication as to three-dimensional geometry. The radiation measured by a white-light coronagraph is photospheric light scattered by free electrons in the highly ionized coronal plasma. Coronal densities are so low that the corona is optically thin in this radiation. Hence the intensity recorded at a given position on a coronal image is an integral of contributions from electrons along the line of sight passing through that position. Different three-dimensional coronal density structures, ranging from "true loops" with small extent along the line of sight to "bubbles" with large extents along the line of sight, can appear to be "loop-like" in a coronagraph image. The actual three-dimensional geometry of loop-like mass ejections remains controversial.

The speeds of the leading edges of mass ejections observed by the Skylab coronagraph (Gosling et al., 1976) ranged from less than  $100$  to greater than  $1200 \text{ km sec}^{-1}$ . The average speed within the  $1.75$  to  $6$  solar radii field of view of that instrument was  $470 \text{ km sec}^{-1}$ . For those ejections seen on enough images to determine the speed at different heights in the corona, all but one showed a nearly constant speed with a small upward acceleration within the field of view.

The masses involved in these events can be estimated by integration of the observed brightness (or polarization brightness) in excess of that in a pre-event image over the field of view (the integral along the line of sight incorporated in the intensity then yields a total mass with only a weak dependence on the unknown three-dimensional geometry). The energy in the event can then



be estimated by assigning a reasonable average speed (such as one-half of the speed of the front of the disturbance) to this material. Such estimates for 24 of the ejections observed with the Skylab coronagraph gave masses in the range  $10^{15}$  to  $2.4 \times 10^{16}$  g., energies in the range  $10^{30}$  to  $1.5 \times 10^{31}$  erg. The average mass was  $5 \times 10^{15}$  g, the average energy  $10^{31}$  erg (Rust and Hildner et al., 1980).

Finally, coronal mass ejections were seen more frequently than anticipated by the observers of the early 1970s (e.g., Hildner et al., 1976). For example, the Skylab coronagraph revealed 77 distinct mass ejections in 217 days of operation (Munro et al., 1979), giving a rate of observation of 0.35 ejections per day. The actual rate of occurrence was undoubtedly higher, perhaps by as much as a factor of two (Hildner et al., 1976), because of corrections for duty cycle of the instrument and for the occurrence of mass ejections away from the solar limbs. The longitude of the limb over which the Skylab mass ejections were seen was well correlated with the locations of active regions or sunspots. Hence it was suggested that mass ejections are physically related to photospheric and chromospheric manifestations of solar activity, and that the rate of occurrence of ejections might vary in phase with the sunspot cycle (Hildner et al., 1976).

An initial step toward understanding the role of coronal mass ejections in solar-terrestrial physics is the examination of the association of specific events with solar and interplanetary phenomena such as flares and shock waves. It was recognized very early in the analysis of the ejections observed by Skylab (Gosling et al., 1974a) that most of these events were not associated with large solar flares or related radio emission and that they occurred more frequently than the well-known interplanetary shock waves. A comprehensive study of associations with photospheric and chromospheric activity (Munro et al., 1979) confirmed the former conclusion. Of the Skylab mass ejections that could be associated, 40% were associated with solar flares, 50% could be associated with eruptive prominences solely (i.e., without any associated flares) and more than 70% with eruptive prominences or filament disappearances with or without flares. Thus Munro et al. (1979) emphasized the physical connection of mass ejections with erupting prominences and a restructuring of the coronal magnetic field in which flares might or might not occur but were not an essential part of the physical evolution of the system. A similar physical interpretation emerged from the study of solar activity seen in X-ray emission. An association between long-duration X-ray brightenings and the loop prominence systems traditionally related to large, two-ribbon flares has been established (e.g., Kreplin et al., 1962; Teske, 1971). Skylab X-ray images (see Moore et al., 1980 and references therein) showed that these long duration X-ray enhancements are produced by arcades of X-ray emitting coronal loops that arched over the previous location of an H $\alpha$  filament that disappeared at the start of the two-ribbon flare. Webb et al. (1976) showed that such X-ray emitting loop prominence systems are characteristic of H $\alpha$  filament disappearances whether or not they occurred in active regions and whether or not they were accompanied by flares. Sheeley et al. (1975) and Kahler (1977) further identified such events with mass ejections seen by the Skylab coronagraph. This again suggests a magnetic restructuring of the outer corona in association with mass ejections, but with the occurrence of the two-ribbon flare as a non-essential (in some cases non-existent) part of the larger-scale physical process. The entire process has been described in terms of magnetic reconnection (see item 3 in Section III).

The disparity in the rates of occurrence of mass ejections and interplanetary shock waves motivated searches for other solar wind features that might be the interplanetary manifestations of coronal mass ejections. As the latter appear as bright, hence abnormally dense, structures in the corona, a logical signature for their interplanetary counterparts might well be a local enhancement of the solar wind density. Abnormally high solar wind densities were known to occur near the fronts of both shock waves and recurrent solar wind streams; these density enhancements

were usually interpreted as the result of interplanetary compression of plasma due to the steepening of a "wave" of fast solar wind. Gosling et al. (1977) identified high density solar wind structures that were not associated with increasing flow speed and thus not subject to local compression. These NCDEs (non-compressive density enhancements) were advocated as the interplanetary counterpart of those mass ejections that did not produce shock waves on the basis of the expectation mentioned above, a crude computation of the expected evolution (Gosling, 1976) or a dense "blob" in the solar wind, and a rough correspondence between the rates of occurrence and masses of the observed mass ejections and NCDEs.

Thus the role of coronal mass ejections in solar-terrestrial physics remained ambiguous throughout the 1970s. Any temptation to regard coronal mass ejections as the "missing coronal link" in the solar flare-interplanetary shock wave-geomagnetic storm chain of cause and effect described above had to be tempered. The evidence indicated that while the missing link might be found among coronal mass ejections, the entire set of these events contained a large population with no obvious place in that chain. Recognition that the most energetic of coronal mass ejections were usually associated (Gosling et al., 1974; Gergely and Kundu, 1974; Stewart et al., 1974; Munro et al., 1979) with the metric radio emission thought to be related to coronal shock phenomena (type II and IV emission) and that only the largest of the masses and energies estimated for the ejections were as large as those in interplanetary shock waves led to suggestions that there were in fact two different classes of coronal mass ejections. One class, produced explosively by solar flares and leading to interplanetary shock waves was the coronal missing link. The other class, perhaps produced by or along with erupting prominences, did not lead to interplanetary shock waves (Gosling et al., 1974a). The interplanetary effects and solar-terrestrial consequences of the latter class remained poorly understood.

### III. THE PRESENT

While the discovery of coronal mass ejections in the early 1970s and the study of their properties and associations suggested intriguing possibilities for their role in solar-terrestrial physics, as many new problems were raised as old problems were solved. In particular, the basic questions as to the solar origins and interplanetary effects of mass ejections remained unanswered or poorly answered. It thus seems appropriate here to summarize the "tools" available for a present-day reexamination of these questions before, in the final section of this report, attempting to set down specific approaches to such a reexamination.

#### THEORETICAL MODELS

The early descriptions of coronal mass ejections stimulated numerous and varied attempts to construct theoretical models to "explain" this phenomenon. The development of such models continues today. Detailed reviews of this material can be found in Dyer (1982), Wu (1982), and Rosner et al. (1983). Our goal here is to briefly describe these models in the context of the first fundamental question stated in the introduction - the origin of coronal mass ejections. That is, can we use models to differentiate between conflicting hypotheses as to the origin of mass ejections by either qualitative or quantitative comparison with observations?

In examining these theoretical models we must not forget that they are attempts to deal with a rather complex physical phenomenon. The very nature of coronal "transients" identifies them as time-dependent phenomena, and any satisfactory model of them must ultimately deal with time-dependent equations of motion. The appearance of mass ejections leaves no doubt that they are two- or three-dimensional spatial features, and any fully satisfactory model of them must deal

with this geometric complexity as well. The generally-accepted view that magnetic effects play a dominant role in the structure of the corona suggests that any realistic model of coronal mass ejections must ultimately deal with magnetic forces, perhaps in the initiation of the ejection, certainly in its interaction with the background or ambient corona. Thus even with simplifying assumptions about the poorly understood thermal structure or energy balance of the corona, a theorist dealing with mass ejections is faced with the description and analysis of a time-dependent, multidimensional, and magnetohydrodynamic physical system.

In response to these inherent difficulties, theorists have concentrated on models with idealized geometries. In most cases, their efforts have been specifically focused on models of "loop-like" mass ejections. Nearly all analyses have been two-dimensional with the plasma and magnetic field of the corona assumed to depend on two spatial coordinates but to be independent (or only weakly dependent) on the third spatial coordinate. These self-imposed limitations must be recognized when any predictions of the models are compared with observations. Restriction of comparisons to loop-like events would appear to present no serious problem since  $\sim 1/3$  of the Skylab mass ejections and  $1/2$  or more of the SMM ejections have a loop-like appearance. Given our poor understanding of the three-dimensional geometry of mass ejections, it is far from clear that they can be represented as two-dimensional objects.

A logical possibility for the origin of coronal mass ejections is the explosive ejection of material by solar flares, as suggested by the solar-geomagnetic and interplanetary shock discussions of Section II. This mechanism has been extensively explored in a series of models by Dryer (1981); Dryer and Maxwell (1979); Dryer et al. (1979);, Maxwell and Dryer (1981, 1982); Nakagawa and Steinolfson (1976); Nakagawa et al. (1975, 1978, 1980, 1981); Steinolfson (1982a, b); Steinolfson and Nakagawa (1976b, 1977); Steinolfson et al. (1977, 1978, 1979, 1981); Wu (1980, 1982); and Wu et al. (1978, 1981a, b, 1982, 1983). A pressure pulse of arbitrary duration (e.g., the result of thermal energy released by the flare) is assumed to occur in a small region at the base of the corona at the initiation of a "transient." The material in this region of high pressure then expands into the surrounding low-pressure ambient corona, with a major expansion toward the outer corona, or in the direction of the greatest pressure difference. This expanding material in turn "sweeps up" the ambient corona to form a compressive shell that moves away from the original site of energy deposition and outward through the corona into interplanetary space. For sufficiently large energy depositions, material is so strongly accelerated that it sweeps into the ambient faster than the characteristic speeds (sonic, Alfvénic, or magnetosonic) in that medium; this compressive wave can then steepen into a shock wave that propagates outward through the corona. This entire physical evolution of the corona, from postulated explosive initiation through the nonlinear atmospheric response, has been followed using two-dimensional (usually in a meridional plane where the independent spatial variables are radius and solar latitude) numerical computations. Descriptions of this work have sometimes invoked specific identification of the resulting, dome-shaped compressive waves (Figure 6-3) with the bright loops observed in coronal mass ejections (Wu et al., 1983; Dryer, 1982; Dryer et al., 1979). The propagation of the shock wave into interplanetary space is then regarded as part of the solar flare-interplanetary shock sequence of events described in Section II.

These models have been extensively developed and described in the literature, within the basic assumptions of a two-dimensional system described by the equations of magnetohydrodynamics. Changes in the coronal response to the explosive initiation have been examined by varying the strength, duration, and location (with respect to the ambient coronal magnetic field structure) of the energy deposition and the state of the ambient corona. Input "thermal functions" deduced from X-ray observations of specific solar flares have been used to derive models of specific coronal mass ejections (Dryer et al., 1979; Dryer and Maxwell, 1979; Steinolfson and Nakagawa, 1977; Wu et al., 1983). Again, the dome-shaped compressive or shock waves predicted in these

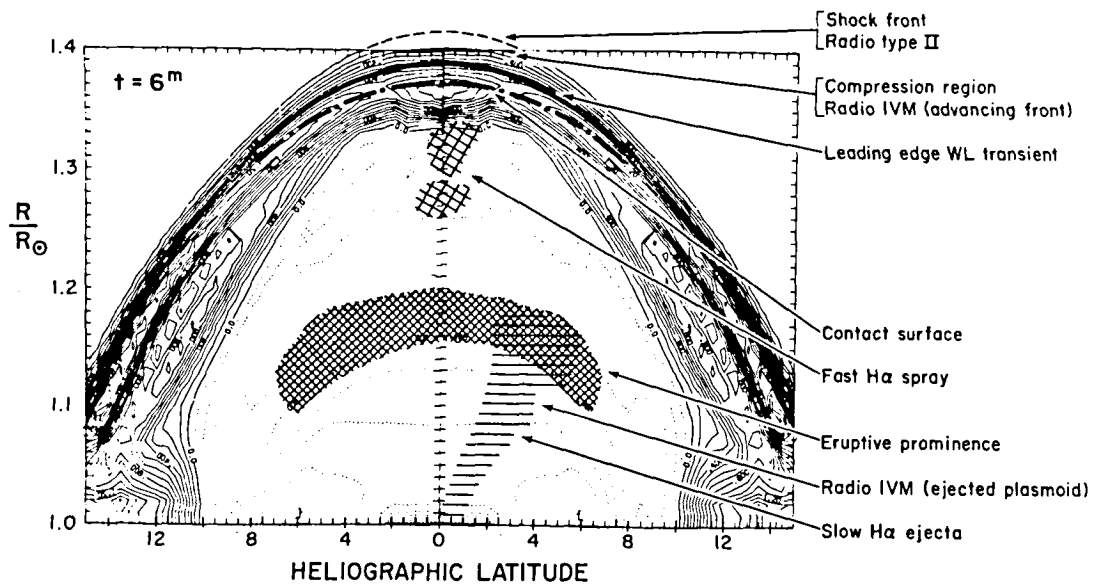


Figure 6-3. A model of flare-associated shock phenomena in the solar corona (Maxwell and Dryer, 19B).

specific examples have been identified with the bright loops observed in the transients and agreement in terms of shape, speed, or general appearance claimed. The models should prove extremely useful in our reexamination of the origin of mass ejections, as they give quantitative predictions of the coronal response to one of the suggestions that must be considered - the solar flare source of coronal mass ejections. As models of compressive and expansive wave propagation through the corona, they should also prove useful in examining the identification of loop-like mass ejections with such waves produced by any localized ejection of material from the low corona.

A number of other suggested origins for coronal mass ejections can be grouped into a very loose class of models unified by not invoking or requiring a flare-like, local release of thermal energy. Rather, these models invoke the forces acting on a flux tube in the coronal magnetic field to drive that structure away from the Sun as part of a coronal mass ejection. Specific mechanisms of this general nature including the following.

- (1) The magnetically driven expansion of a "rope" or loop of magnetic field lines (Mouschovia and Poland, 1978; Anzer, 1978).
- (2) The buoyant uplift of a magnetic loop immersed in the inhomogeneous ambient coronal pressure and magnetic field (Yeh and Dryer, 1981a, b; Yeh, 1982, 1983).
- (3) The magnetically driven uplift of flux tubes at the top of a magnetic arcade, produced by a gradient in magnetic pressure large enough to overcome the downward magnetic tension and gravity (Pneuman, 1980a, b). The magnetic configuration necessary for this mechanism is advocated on the basis of reconnection of a distended magnetic arcade beneath the transient as suggested by the X-ray observations summarized above in the section entitled "Coronal Transients or Mass Ejections."

- (4) The expulsion of diamagnetic plasmoids (regions with closed magnetic field lines disconnected from the Sun) by a net upward magnetic force due to the gradient of magnetic pressure in the ambient corona (Pneuman, 1983).
- (5) An abrupt departure from a magnetostatic equilibrium state with closed magnetic loops, attained during a slow change in boundary conditions (Low, 1981).

In all descriptions of these models, the bright features observed in loop-like coronal mass ejections are identified with the magnetic structures inherent in the mechanisms -- magnetic loops in 1-2, the rising tops of arcades or plasmoids in 3 and 4, and the region of rising closed field lines in 5.

All of these mechanisms (1-5) have been shown to be plausible on the basis of stability, force balance, or quasi-equilibrium analyses of highly idealized magnetic structures. In no case has a full, self-consistent integration of the time-dependent MHD equations confirmed the validity of the mechanism without rather arbitrary assumptions about the magnetic geometry and its temporal evolution. In no case has the interaction of the proposed expanding structure with the ambient corona been taken into account. Low (1981a, b) has extended the concept of a structured ejecta as the features observed in mass ejections by finding self-similar solutions to the equations of magnetohydrodynamics and using them to describe the motion of the ejecta through the corona. These models are then far less developed and quantitative than the flare-compressive wave class described earlier. Any comparison with observations of mass ejections must be qualitative or deal with general quantitative properties such as the speed or shape of loop-like features.

We thus have a mixed bag of models for use in examining the origin of coronal mass ejections. Two different classes of models postulating flare and non-flare origins are available, providing some hope for distinguishing between these two natural hypotheses. Given the very different states of development of these two classes, any comparisons with observations will be at very different levels of detail; we must resist the temptation to regard any failure of, for example, the flare-shock models to match mass ejection observations in detail as an argument for one of the other models. The two classes also invoke different identifications of the bright loop-like features in some mass ejections, propagating compressive or shock waves (and perhaps expansive waves) in contrast to magnetic loops or arcades frozen into the expanding atmosphere. We thus have another important difference in interpretation that may be tested by comparison with observations.

## NEW OBSERVATIONS OF CORONAL MASS EJECTIONS

We have been fortunate to have available at the time of this workshop several new sets of data pertaining to coronal mass ejections and their role in solar-terrestrial physics. Of great significance to this discussion are three new coronagraph systems that extend, both upward and downward, the range of coronal heights in which mass ejections can be observed, that have time resolution superior to that previously available, and that add coronal images in two important emission lines to the white-light data upon which earlier studies were based. The SOLWIND white-light coronagraph has been monitoring the corona from the satellite P78-1 since March, 1979 (Sheeley et al., 1980). The coronal field of view of this instrument extends from 2.5 to 10 solar radii; images are obtained at intervals as short as 5 minutes during the one hour daylight portion of each 97 minute spacecraft orbit. The coronagraph/polarimeter flown on the Solar Maximum Mission operated between March and September, 1980 (Csoeke-Poeckh et al., 1982; House et al., 1981). Its field of view was between 1.6 and 6 solar radii, with images through clear and

polaroid filters obtained at intervals as short as 3 minutes during the 50-minute daylight portions of 96 minute orbits. In addition to white-light observations, this instrument obtained images with filters that isolate H $\alpha$  and the coronal green line (5303Å) emission. Finally, a new ground-based coronameter system has been operated at the Mauna Loa Solar Observatory since early 1980 (Fisher et al., 1981; Fisher and Poland, 1981). It measures the coronal polarization brightness between 1.2 and 2.2 solar radii with a time resolution of  $\sim 10$  minutes and has proven capable of detecting mass ejections in the differences between successive images. The existence of these three coronagraph systems offers the possibility for studying the properties of coronal mass ejections over the wide range of heliocentric distances between 1.2 and 10 solar radii with high (a few to ten minutes) time resolution.

Complementary solar and interplanetary observations provide further opportunities to test and quantify suggestions concerning the relationship of coronal mass ejections to solar activity and interplanetary phenomena. The usefulness of X-ray observations in understanding the "activity" related to Skylab mass ejections has been summarized in Section II above. X-ray and UV observations from instruments on the SMM and GOES spacecraft provide similar information for the recent era. Radio observations at long wavelengths ( $10^2$  to  $10^4$ m) made from the ISEE-3 spacecraft (Cane et al., 1982) complement ground-based observations and make it possible to trace type II radio emission over heliocentric distances from 10 solar radii to the orbit of Earth. Penetration of the Helios spacecraft into the inner solar system has led not only to a better understanding of the evolution of solar wind structure with heliocentric distance, but to observations of solar wind properties over the solar limbs where nearly all coronal mass ejections have been detected.

The new observations of coronal mass ejections are still in the early stages of analysis and comparison with solar and interplanetary data. Any attempt to comprehensively review this material would be premature. Nonetheless, these ongoing studies have already yielded factual information about coronal mass ejections that confirms, refines, or confounds our earlier understanding of this topic. These recent results, despite their newness or even preliminary nature, were an important part of our discussions at this workshop and must be integrated into any overview of coronal mass ejections.

### 1. Comparison of Mass Ejections seen by Skylab, SMM, and SOLWIND Coronagraphs

Comparisons of the mass ejections seen on the declining phase of sunspot cycle 20 (Skylab) and near the maximum of sunspot cycle 21 (SMM and SOLWIND) show, as expected, the rate of occurrence of major mass ejections is higher near sunspot maximum. However, the rate derived from the SMM observations was only slightly higher (Hundhausen et al., 1983) than the duty-cycle corrected rate from Skylab; the change was less than expected on the basis of a simple linear relationship to sunspot number (Hildner et al., 1976). The SOLWIND observations show somewhat higher rates than SMM for longer intervals near sunspot maximum (Howard, 1983). Both the SMM and SOLWIND coronagraphs have seen mass ejections at high solar latitudes during the sunspot maximum epoch (Hundhausen et al., 1983; Sheeley et al., 1982b). No Skylab mass ejections were centered at position angles more than  $\sim 50^\circ$  from the solar equator (Hildner, 1977); in contrast 27% of the SMM mass ejections were centered more than  $45^\circ$  from the equator. The H $\alpha$  images obtained by the SMM coronagraph confirm that most of the material seen in loop-like transients is coronal (not emitting in H $\alpha$  and thus containing little neutral hydrogen). However, H $\alpha$  emitting material is seen to transit the entire field of view in some prominence-associated mass ejections (House et al., 1981). This material often has the appearance of twisted ropes, much like erupting prominences observed in the low corona.

The surprises arising from these comparisons come for those who might have expected many more mass ejections, more flare-associated ejections or faster mass ejections at a time of high

solar activity. The associations of mass ejections with photospheric and surface activity during the SMM epoch (Sawyer et al., 1983) shows a pattern very similar to that found for the Skylab epoch. More mass ejections can be reasonably associated with erupting prominences than with flares. Sheeley et al. (1983a) have found a very high percentage of associations of mass ejections with long-duration 1-8Å X-ray enhancements observed from the GOES satellite. Preliminary values of the speeds derived for the mass ejections seen by the SOLWIND and SMM coronagraphs show a distribution not much different from that obtained during the Skylab mission. If any change in the distribution can be seen, it is in the direction of more low speed events near sunspot minimum.

## 2. Mass Ejections in the Low Corona

The ground-based coronameter at the Mauna Loa Solar Observatory has detected numerous transient changes in the low corona. Some of these correspond to mass ejections seen in the SMM field of view. In the low corona these events have the appearance of mound-shaped reductions of coronal polarization brightness, surrounded by a diffuse shell of enhanced polarization brightness. An unexpected result of the observations has been the detection of a common class of "dark" coronal transients in which only the mound-shaped region of reduced brightness is visible in the early stages of the event (Fisher and Poland, 1981). Low et al. (1982) have suggested that these density-depleted regions are permeated by closed magnetic field lines, expanding rapidly outward as their equilibrium with the surrounding atmosphere approaches the abrupt change described under mechanism 5 in the first part of this section.

MacQueen and Fisher (1983) have used the Mauna Loa coronameter observations in a study of mass ejection speeds as a function of height in the low corona. They found a clear delineation between the dynamical behavior of flare-associated and eruptive prominence-associated ejections as displayed in Figure 6-4. The flare-associated events had high (greater than 300 km sec<sup>-1</sup>) nearly constant speeds. In contrast, erupting prominence-associated ejections were all moving at lower speeds (~50 to 100 km sec<sup>-1</sup>) when first detected and were significantly accelerated through the 1.2 to 2 solar radii height range accessible to this instrument; they sometimes attained speeds comparable to the flare-associated ejections before leaving the field of view. This seems to be the first observation of a consistently different dynamical behavior, suggesting an acceleration mechanism different in nature or in time scale, for ejections with different associated activity in the lower solar atmosphere. It is a strong argument for the existence of two different physical classes of mass ejections as suggested in Section II. The history of solar-terrestrial physics bears witness to the folly of seeking a single cause for two physically distinct phenomena (e.g., recurrent and sporadic geomagnetic activity), and pursuing a single cause (such as sunspots) long after it should have been recognized that they could not be the source of one (ultimately both) of these classes. We must guard against making a similar mistake if we fail to consider two classes of coronal mass ejections and continue attempts to describe both as part of the flare-shock wave phenomenon.

## 3. Mass Ejections in the Outer Corona

The SOLWIND coronagraph has detected a large number of transient changes in the outer corona at heights of 2.5 to 10 solar radii (Sheeley et al., 1980). These observations have proven to be particularly useful in the study of associations with interplanetary phenomena as will be described below. They have also been used, in the manner mentioned above, to extend the range over which coronal mass ejections can be followed and their properties determined. For example, Wagner et al. (1983) have examined a prominence-associated, loop-like mass ejection on April 15-16, 1980 seen by both the SMM and SOLWIND coronagraphs. The height of the white-light loop as a function of time is shown on Figure 6-5; SMM positions are in the range

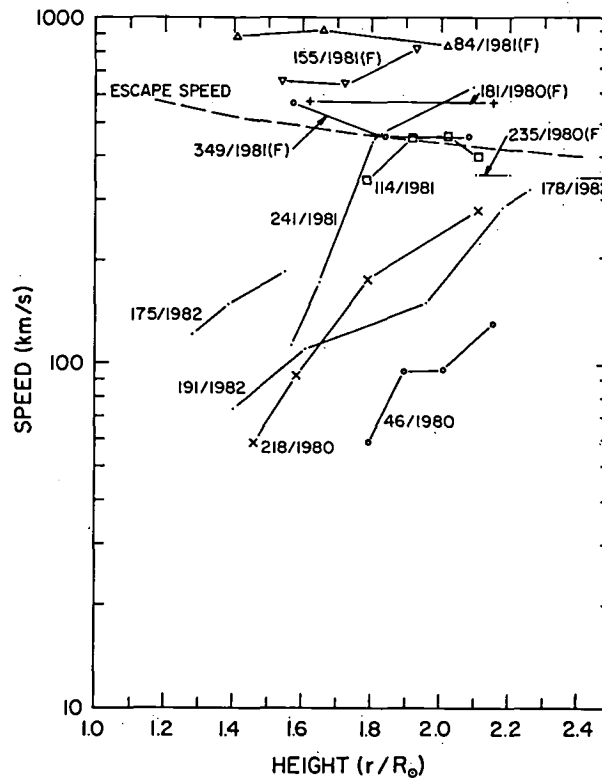


Figure 6-4. The radial speeds of coronal mass ejections as a function of height in the low corona (MacQueen and Fisher, 1983). The flare-associated ejections are identified by the (F) symbol following the date for the event.

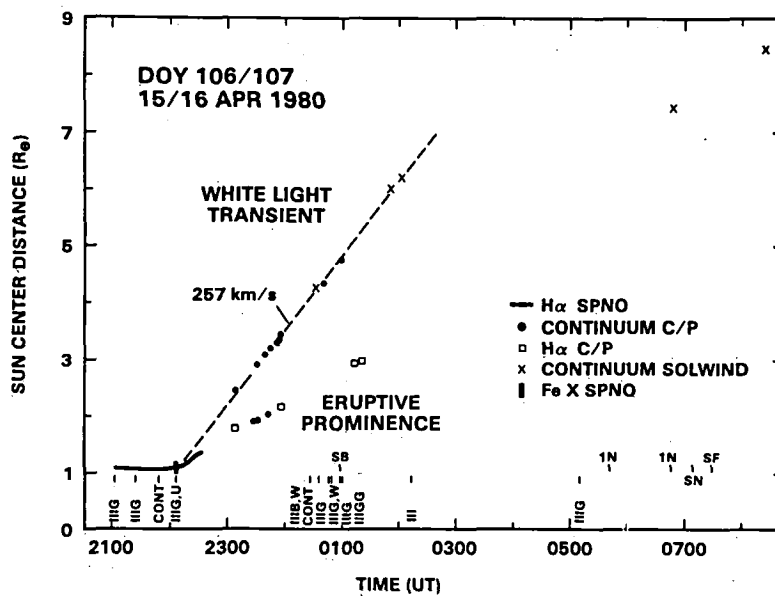


Figure 6-5. The position (heliocentric distance in solar radii) vs. time plot for the white loop and eruptive prominence observed in the coronal mass ejection of April 15-16, 1980 (Wagner et al., 1983).



2.5 to 5 solar radii, SOLWIND positions in the range 4.3 to 6.2 solar radii. A single, straight-line height vs. time relation fits all positions well, indicating that the loop moved at a nearly constant speed of  $257 \text{ km sec}^{-1}$  over this entire height range. This line also intersects the base of the corona at a time within a few minutes of the eruption of the associated prominence. The height vs. time curve for  $\text{H}\alpha$  emitting material seen by the SMM instrument extrapolates to a similar intersection with the base of the corona; this curve again implies motion at a nearly constant speed, but with a lower value,  $205 \text{ km sec}^{-1}$ . Two observations of a white-light feature at 7 to 9 solar radii by the SOLWIND instrument fall near a straight line drawn through the  $\text{H}\alpha$  positions, suggesting that this material may have been detected in the continuum at these great coronal heights. Other events lead to a similar suggestion;  $\text{H}\alpha$  emission continues as material from an eruptive prominence moves well out in the corona and this physical structure may retain its identity, observable because of its high density even in the white-light continuum, throughout the SOLWIND field of view.

#### 4. Associations Between Coronal Mass Ejections and Interplanetary Phenomena

Early attempts to identify the interplanetary counterparts to coronal mass ejections were based on the main distinguishing features of these events as seen in the corona, the transient addition of mass and energy to the ambient coronal expansion. Two types of solar wind structures showed these same characteristics. The first was the interplanetary shock wave introduced in Section II; the second was the class of solar wind density enhancements that were not readily identified as caused by a compression of the ambient flow, the so called non-compressive density enhancements or NCDEs. Estimates of the extra mass and energy fluxes carried by such structures (Hundhausen et al., 1970; Gosling et al., 1977) yielded values of the order of those estimated from the Skylab coronagraph images mentioned previously. Although the initial suggested associations were based on indirect evidence, more recent studies have placed these identifications on a firmer footing. The initial direct observation of a single flare-related mass ejection and its interplanetary shock wave disturbance was reported by Gosling et al. (1975). During the long operation of the SOLWIND coronagraph, the Helios spacecraft passed through aphelion over the solar limbs (as viewed from the Earth) offering an unprecedented opportunity for the direct association of observed coronal phenomena with interplanetary phenomena observed at similar solar longitudes. Comparison of the two data sets (Schwenn, 1983; Sheeley et al., 1982a, 1983b) reveals that nearly all of the low-latitude, high-speed coronal mass ejections observed by SOLWIND were related to shock waves seen at Helios. Figure 6-6 shows an example of a mass ejection seen in the coronagraph images and the variations in plasma parameters signaling the arrival of a shock wave and perhaps the driver gas or ejecta at the 0.43 AU location of Helios some 30 hours later. The speeds of the coronal mass ejection, the interplanetary shock wave, and the average transit speed implied by the association tend to fit the view that interplanetary shock waves are gradually decelerated in passage from the corona through the solar wind. Some Helios shocks were associated with low speed (less than  $400 \text{ km sec}^{-1}$ ) suitably directed mass ejections suggesting either acceleration of the shock beyond 10 solar radii or that the coronal mass ejection is not the shock wave. This same study also identified several NCDEs as the interplanetary counterparts of coronal mass ejections. However, it is now believed that the majority of NCDEs are the interplanetary manifestations of coronal streamers (Borrini et al., 1981; Gosling et al., 1981; Feldman et al., 1981) rather than CMEs. However, about 25% of the detections of heavy ion spectra, indicating unusually hot coronal conditions, have been interpreted to indicate that at least some NCDEs have a transient coronal origin (Fenimore, 1980).

Analysis of solar wind observations have suggested characteristics of post-shock solar wind flows identify the driver gas mentioned in Section II (Burlaga et al., 1981; Zwickl et al., 1983; Neugebauer, 1983). Among these are enhancements in the helium abundance, anomalously low electron and proton temperatures, high ionization state temperatures, high magnetic field

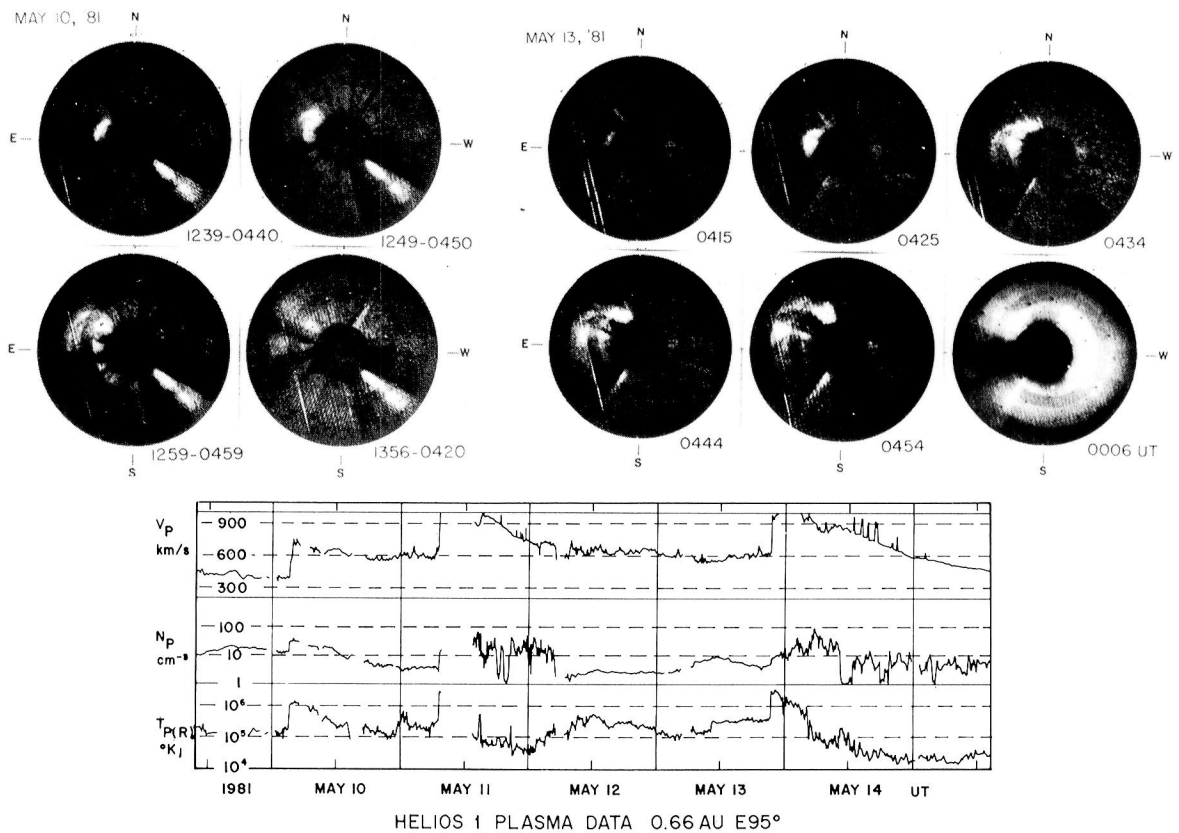


Figure 6-6. SOLWIND coronagraph images showing coronal mass ejections on May 10 and May 13, 1981 and Helios plasma data, showing shock waves on May 11 and 13 that have been associated (Sheeley et al., 1983) with the ejections.

strengths, bidirectional streaming of suprathermal electrons and energetic ions, low variances in the bulk speed and magnetic field, highly structured densities, and gross rotations in the interplanetary magnetic field designated as "magnetic clouds". An operational complication of any such identification is that each of these identifiers, as well as the many different possible combinations of the various members of this list, are neither always present nor mutually overlapping after the occurrence of an interplanetary shock. Furthermore, these same solar wind characteristics have also been observed when no shock has been detected (see e.g., Zwickl et al., 1983).

Several of the foregoing characteristics tentatively identified with driver gas place constraints on models describing the physical mechanisms which contribute to coronal mass ejections. Most important are measurements of anomalous ionization states of coronal heavy ions embedded in the flows (Bame et al., 1979; Fenimore, 1980; Schwenn et al., 1980; Gosling et al., 1980; Zwickl et al., 1982). Although most of these measurements are consistent with ionization temperatures in the range between  $2.5$  and  $3.5 \times 10^6 \text{ K}$ , temperatures as high as  $15 \times 10^6 \text{ K}$  and as low as those normally found in the chromosphere were substantial abundances of  $\text{He}^+$  are present have been found. Whereas the high and low ends of this range are consistent with those measured in the hot kernels of solar flares (Vaiana and Rosner, 1980) and in the  $\text{H}\alpha$  prominences observed to be ejected in association with some coronal mass ejections (see the first part of this section), the midrange of temperatures indicated by most heavy ions spectra are consistent with average active region coronal temperatures (Vaiana and Rosner, 1980).

A second important characteristic tentatively identified with the driver gas and bearing on the mechanism of coronal mass ejection is that the ejected material is usually at a higher pressure

than the neighboring gas and may thus be expanding at 1 AU (Hirshberg and Colburn, 1969; Schatten and Schatten, 1972; Burlaga and King, 1979; Borrini et al., 1982a, b). Furthermore, this overpressure is dominated by the magnetic field terms  $B^2/8\pi$  which is higher and less variable than normal. In contrast the density as well as the He abundance is highly variable. Because reasonable scaling laws indicate that the ratio of plasma to magnetic field pressure in the expanding solar wind should be lower near the Sun than near 1 AU, these results imply that the magnetic field most likely dominates the pressure balance of mass ejections in the low corona and would allow large fine-scale spatial variations in the plasma internal state. However, the observed large variances in the density and chemical composition at 1 AU suggest evolution from a highly structured corona which was either pre-existing or which resulted in response to the transient mass ejection. Such structure might be identified with the many nested bright loops so evident in coronagraph images of coronal mass ejections.

A last feature of driver gas bearing on the mechanism of mass ejection is that it often contains bi-streaming suprathermal electrons and energetic ions. In addition, the temperatures of both the thermal electrons and ions are often anomalously cold. These observations have been interpreted to indicate a closed-loop magnetic topology (Gosling et al., 1974b; Montgomery et al., 1974; Palmer et al., 1978; Pudovkin et al., 1979; Temny and Vaisberg, 1979; Bame et al., 1981; Kutchko et al., 1982). Such a topology is consistent with the coherent rotations of the magnetic field observed sometime within driver gas and given the name of magnetic clouds (Burlaga et al., 1981; Klein and Burlaga, 1982; Burlaga and Behannon, 1982). It suggests a resolution of the question regarding the magnetic topology of the flare ejecta, raised in Section II, in favor of the "pinching-off" of distended magnetic loops. It should be pointed out, however, that conclusive proof of the existence of such magnetic structures has yet to come from direct observations of the interplanetary magnetic field.

The SOLWIND coronagraph has also observed a new type of coronal mass ejection with extremely interesting consequences for the study of interplanetary and geomagnetic effects. Subtraction of coronagraph images sometimes reveals a "halo" of enhanced brightness gradually expanding about the Sun (Howard et al., 1982). This apparition has been interpreted as a mass ejection moving, with a broad front, directly, toward the Earth, as shown on Figure 6-7. For one such event seen on November 27, 1979, both a disappearing filament and a small solar flare that occurred between 0500 and 0700 UT could be associated with this event; an interplanetary shock wave was detected on November 30, 1979 at 0649 UT by the ISEE-3 spacecraft and a storm sudden commencement began at the Earth at 0738 UT on November 30, 1979. If the mass ejection is assumed to be a spherical shell, the observed rate of expansion of the observed Sun-centered disc,  $600 \text{ km sec}^{-1}$  implies a frontal radial expansion at  $1160 \text{ km sec}^{-1}$  with the front of the sphere, in the direction of the Earth, at a heliocentric distance near 10 solar radii. Comparison of the inferred mass ejection speed of  $\sim 1000 \text{ km sec}^{-1}$ , the speed implied by transit time to the orbit of Earth,  $\sim 560 \text{ km sec}^{-1}$ , and the local speed of the shock front near the Earth,  $\sim 400 \text{ km sec}^{-1}$ , again suggests a slow deceleration of the shock and ejecta in the interplanetary region. The very appearance of this feature implies a broad, bubble-like rather than loop-like ejection in the outer corona.

## 5. Associations Between Coronal Mass Ejections and the Phenomena Revealed by Radio Observations

Recent studies (communicated at the workshop by R. T. Stewart and R. G. Stone) comparing meter wavelength type II bursts and coronal transients have shown several important results:

- The starting frequencies of meter type II bursts associated with coronal mass ejections are substantially lower than type IIs without the association. The implication is that the starting height of the shock is higher for the coronal mass ejection events.

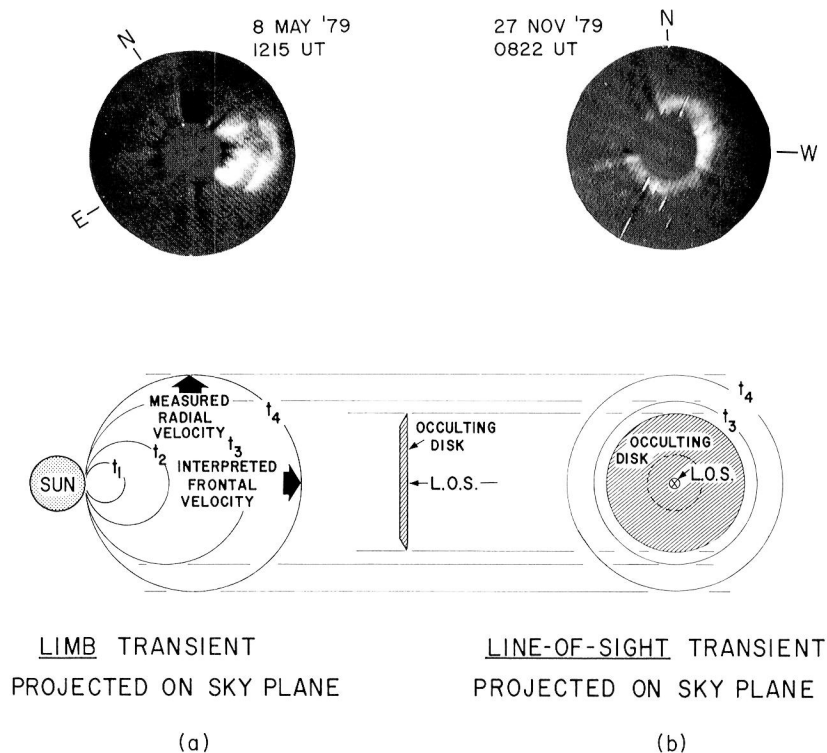


Figure 6-7. Coronal mass ejections seen over the solar limb (left) and moving directly toward the Earth (right) (Howard et al., 1982).

- Simultaneous radio and white-light observations in the lower corona for the four available events suggest that the type II emission is not coming from the leading edge of the transient as has been proposed for earlier models (Gary et al., 1982; Gergerly et al., 1983a, b). Whether propagation effects can account for this result is not clear at the present time. However, recent interplanetary type II observations show that the radio emission occurs from behind the shock.
- Both meter and kilometer type II bursts which accompany coronal mass events are associated with long duration soft X-ray events lasting  $\sim 3$  hours whereas type II bursts without coronal mass ejections are associated with much shorter duration X-ray events lasting  $\sim 1/2$  hour (Sheeley et al., 1983a). The implication here is that coronal mass ejections and some type II bursts are associated with a sustained and gradual input of thermal energy low in the corona which could account for the shock driver.

#### IV. THE FUTURE

Given the background material in Section II and the basic questions concerning the solar origins and interplanetary effects of coronal mass ejections, let us conclude with a discussion of approaches, both broad and specific, to these questions using the tools described in Section III.

## SOLAR ORIGINS

The theoretical models described in Section III are, of course, attempts to understand the response of the solar corona to changing physical conditions near its lower boundary. They can help us to understand the solar origins of coronal mass ejections in three different ways (or at three different levels).

1. As aids in understanding the basic physics of initiation and propagation of disturbances in this complex physical system and thus as a source of "educated intuition" concerning its behavior.
2. As broad or general models of a class coronal mass ejection that can be compared qualitatively with observations to sharpen the "educated intuition" mentioned above and yield some sense of the plausibility or validity of the assumptions made in the models.
3. As detailed models of some coronal mass ejections that can be compared qualitatively with observations and yield stronger tests of validity of assumptions, serving as guides to their revision or refinement.

All of the models described in Section III have been developed to the point at which they can be used in the first two ways listed above. The two-dimensional numerical models of flare-initiated coronal compression or shock waves and some of the other models described in Section III have also been used to simulate specific mass ejections and thus regarded as entering the third class of utility.

It became very clear in the discussions at this workshop that our "intuition" concerning the propagation of disturbances in the solar corona has not been well developed or educated through use of existing theoretical models. In all too many cases, specific models have been refined to levels of sophistication where they are proposed and used at the second and third level in our list without adequate exploration of basic physical understanding and transmission of that knowledge to the entire community. Stated directly, it seems to be more fashionable for both theorists and observers to consider qualitative comparisons that seem to establish the "rightness" of a model than to probe for general, basic understanding. Even the most familiar and sophisticated of the theoretical models of coronal mass ejections can and should continue to be used through parameter studies and careful analysis of their predictions at the first level in our list.

On the other hand, any use of theoretical models at the third level in our list must, at this time, be made with great caution. For example, the two-dimensional numerical computations of flare-initiated shock waves can be carried out in great detail and have been proposed as simulations of specific coronal mass ejections. However, it is well understood that the results of these computations depend, sometimes sensitively, upon the inferred or assumed flare input signal and upon the poorly known and hence highly idealized simplification of the ambient corona through which the shock propagates. Both the proposed simulation of events by such models and any comparisons with observations must be made with this source of uncertainty specifically considered and stated. Any comparison, qualitative or quantitative, between the models and the observations must involve quantities that are essential features or predictions of the models and are actually measured or inferred from the observations. If models are proposed as quantitative and compared with observations at this third level, some standard for an acceptable level of agreement must be found. Very little consensus as to that level was arrived at in our discussions.

Given these rather philosophical statements concerning the use of theoretical models and their comparison with observations, what are the practical realities? Most of our empirical knowledge of coronal mass ejections comes from their observation with white-light coronagraphs. These instruments yield a two-dimensional image of the corona, hopefully at a sequence of closely spaced times. Thus we have, at best, a statement of the white-light coronal brightness  $I$  as a function of distance from the center of the Sun,  $r$ , position angle around the limb of the Sun,  $\theta$ , and time  $t$ , over a range of  $r$  defined by the field of view of the instrument and a range of  $t$  determined by the operation of the instrument;  $I(r, \theta, t)$ . Some instruments provide (see Section III) a similar determination of the polarization of the white-light corona and of similar quantities integrated over one or two emission lines. It must be emphasized that this is a rather sparse source of information; the common white-light observations are related to a single physical parameter, the coronal electron density, through a line of sight integral (or two integrals if polarization measurements are available). The emission line observations add information on temperature and chemical composition, again integrated along the line of sight. There is no direct information on the properties of the coronal magnetic field. Observations of metric radio emission yield some information concerning the mass ejection (e.g., the presence of and motion of a shock wave) but may still suffer from problems of interpretation that limit their utility in direct comparison with models.

How do we then compare theoretical models with these observations? The latter, through the  $I(r, \theta, t)$  described above, do describe the shape, size, and temporal evolution of a mass ejection for comparison with the density structure predicted by a model. The speed vs. height comparisons that have often been offered as the touchstone of validity are thus but a small part of the comparison that can be made. It seems obvious that the full comparison with  $I(r, \theta, t)$  should be considered and that this will remain the primary observed quantity for the near future. In the past, the predicted evolution of the coronal density structure,  $n_e(r, \theta, \phi, t)$  has been used in qualitative comparisons. There was agreement between the theorists and most observers at this workshop that, until the three-dimensional shape of mass ejections is determined from measurements and the observed  $I(r, \theta, t)$  can be deconvolved to yield an observed  $n_e(r, \theta, \phi, t)$ , it would be more appropriate to compare an integrated  $n_e(r, \theta, \phi, t)$  in a geometry appropriate to the models with the observed  $I(r, \theta, t)$  for any future comparisons.

One "topological" identification within the observed  $I(r, \theta, t)$  structure would seem to promise considerable clarification in using the models as evidence for the origin of mass ejections. If a region of the low corona is ejected outward, by the release of energy in models of flare association or by the outward driving of a magnetic structure in the other models of Section III, a contact surface (or tangential discontinuity) will separate the ejected material from ambient coronal plasma swept up by the ejecta. The flare-shock class of mass ejection models described in Section II deal mainly with this sweeping up of ambient plasma and identifies the bright features of loop-like mass ejections with the shell of compressed plasma. Unless detailed information as to the nature of the energy deposition and the structure of the atmosphere near the deposition site is available and used in such models, they should not be expected to describe the structure well behind the contact surface. In contrast, all of the other models described in Section II deal with the acceleration of the ejecta and do not take into account the sweeping-up of the ambient atmosphere. They thus apply to the structure well behind the contact surface and make no predictions as to structure near it. Identification of the interface between ejected and swept-up material in an observed mass ejection would thus be an important step in assessing the relative importance of the structure inherent in the ejecta and that of the compressed shell in the appearance of coronal mass ejections. This identification would be a guide toward proper comparisons between models and observations. While many observers are tempted to regard the complex, loop-like features seen within mass ejections as evidence for a magnetic structure within the ejecta, this identification is not considered to be proven. The careful analysis of mass ejections observations and the association with solar radio observations (see Woo et al., 1982) might be useful in pursuing this identification.

Finally, it is widely recognized that the study of coronal mass ejections remains in the stage described by Chree in the Introduction, where "...minute study of the facts..." remains the key to understanding. The new body of coronal and interplanetary observations summarized in Section III will, in themselves, undoubtedly reveal many clues as to the origins of mass ejections. For example, the associations of mass ejections with solar flares, erupting prominences, X-ray and radio emission will continue to be a primary source of evidence for the origins of these events. Some knowledge of the structure of the ambient corona into which a mass ejection moves may be helpful in understanding the importance of the interaction between mass ejections and the ambient medium. The detailed temporal behavior of mass ejections may favor or rule out initiations on different temporal or spatial scales. Examination of this behavior might lead to classification of mass ejections and discrimination among various mechanisms for their initiation (as in MacQueen and Fisher, 1983). One long-standing problem in this study of mass ejections, their three-dimensional geometry, looms larger as we view the future. Evidence continues to accumulate for both loop (thin) and bubble (thick) structures in mass ejections. Until this question is settled by observers, theorists attempting to model this coronal phenomenon (or phenomena) lack the basic understanding of its "dimensionality" and all comparisons between models and observations will remain uncertain.

## INTERPLANETARY EFFECTS

Several studies of solar wind data have the potential of both improving our understanding of the interplanetary effects of coronal mass ejections and yielding information on the physical mechanisms of the ejections themselves. Some of these topics can be pursued using existing data bases but others require development of new experimental opportunities using state-of-the-art instrumentation.

Magnetic Topology of Driver Gas. Coronal images of mass ejection events show nested loops of bright material moving away from the Sun. It is not clear at present whether these loops represent density-loaded magnetic field lines or shock fronts. If they outline magnetic field lines, as often assumed, the role of magnetic reconnection before, during and after "lift off" reappears as an important question. The bearing of this question on the mechanism driving the mass ejection suggests considerable attention in future research; see, for example Illing and Hundhausen (1983). Quantitative analyses of existing solar wind and energetic ion data may provide some useful clues in this regard.

Heavy-Ion Spectroscopy. The ionization state of the solar wind provides information about electron temperatures in the corona between about 1.1 and ~6 solar radii. Any such information bears on the processes that expel coronal plasma. Interpretation of these data would be facilitated by analysis of simultaneously observed soft X-ray images of coronal regions in the neighborhood of disappearing filaments.

Theoretical Models of Chemical Fractionation Mechanisms. A characteristic feature of solar wind plasma associated with the shock-wave driver gas and hence some coronal mass ejections is enhanced yet variable  $\text{He}^4$  abundance. Although detailed analyses have not yet been completed, a large data base of  $\text{He}^3$  abundances measured using the Goddard mass spectrometer aboard ISEE-3 is available to investigate whether  $\text{He}^3$  varies in a manner similar to  $\text{He}^4$ . Regardless of the outcome of such a study, the causes of the known, strong  $\text{He}^4$  variations must bear on conditions in, and/or dynamical processes operating within the solar atmosphere near the sites of coronal mass ejections. However, very little is presently known about the mechanisms which can cause chemical fractionation in such regions. This is a wide open field of theoretical endeavor and any information could help organize a large body of data that may prove relevant to the physical mechanisms leading to mass ejections.

Correlative Studies of Coronal Ejections and Their Interplanetary Effects. A fundamental uncertainty in the study of coronal mass ejections is their three-dimensional geometry. Neither 2-D measurements such as provided by coronagraph and X-ray images nor 1-D measurements such as provided by in situ solar wind measurements have resolved this problem. However, a combination of pertinent coronal white-light and X-ray images, type II radio emission intensities, satellite-source radio scintillation measurements, and multipoint in situ measurements are available. A coordinated analysis of these data could yield much detailed information concerning the nature and spatial structure of the driver gas, the spatial relationship between the bright loops so prominent in coronagraph images and a possible shock front near the corona, the 3-D global extent of the resulting interplanetary disturbances and the effectiveness of coronal mass ejections in generating interplanetary shocks.

Type II Generation Mechanism and Their Spatial Extent. Before type II radio emissions can be used as a sensitive quantitative tool for exploring coronal mass ejection mechanisms and their interplanetary effects, it is first necessary to understand the physical mechanisms which convert a plasma disturbance to radio waves. Although there is overwhelming evidence to indicate that type II emissions come from shock wave disturbances, the relationship known from in situ measurements is not one-to-one. However, recent advances in our understanding of the processes which lead to electron heating and wave generation in collisionless shocks lead us to believe that a concerted effort to understand type II emission mechanisms using ISEE-3 data should lead to substantial progress.

## V. REFERENCES

- Akasofu, S. -I., and S. Yoshida, The structure of the solar plasma flow generated by solar flares, Planet. Space Sci., **15**, 39, 1967.
- Anzer, U., Can coronal loop transients be driven magnetically? Solar Phys., **57**, 111, 1978.
- Bame, S. J., J. R. Asbridge, W. C. Feldman, E. E. Fenimore, and J. T. Gosling, Solar wind heavy ions from flare-heated coronal plasma, Solar Phys., **62**, 179, 1979.
- Bame, S. J., J. R. Asbridge, W. C. Feldman, J. T. Gosling, and R. D. Zwickl, Bi-directional streaming of solar wind electrons  $> 80$  ev: ISEE evidence for closed field structure within the driver gas of an interplanetary shock, Geophys. Res. Lett., **8**, 173, 1981.
- Borrini, G., J. M. Wilcox, J. T. Gosling, S. J. Bame, and W. C. Feldman, Solar wind helium and hydrogen structure near the heliospheric current sheet: a signal of coronal streamers at 1 AU, J. Geophys. Res., **86**, 4565, 1981.
- Borrini, G., J. T. Gosling, S. J. Bame, and W. C. Feldman, An analysis of shock wave disturbances observed at 1 AU from 1971 through 1978, J. Geophys. Res., **87**, 4365, 1982a.
- Borrini, G., J. T. Gosling, S. J. Bame, and W. C. Feldman, Helium abundance enhancements in the solar wind, J. Geophys. Res., **87**, 7370, 1982b.
- Burlaga, L. F., and J. H. King, Intense interplanetary magnetic fields observed by spacecraft during 1963-1975, J. Geophys. Res., **84**, 6633, 1979.



- Burlaga, L. F., E. Sittler, F. Mariani, and R. Schwenn, Magnetic loop behind an interplanetary shock: Voyager, Helios, and Imp B observations, J. Geophys. Res., **86**, 6673, 1981.
- Burlaga, L. F., and K. W. Behannon, Magnetic Clouds: Voyager observations between 2 and 4 AU, Solar Phys., **81**, 18, 1982.
- Cane, H. V., R. G. Stone, J. Fainberg, R. T. Stewart, J. L. Steinberg, and S. Hoang, Type II solar radio events observed in the interplanetary medium, Solar Phys., **78**, 187, 1982.
- Carrington, R. C., Description of a singular appearance seen in the sun on September 1, 1859, Mon. Not. R. Astron. Soc., **20**, 13, 1860.
- Chapman, S., and J. Bartels, Geomagnetism, Oxford University Press, London, 1940.
- Chree, C., Studies in Terrestrial Magnetism, MacMillan and Co., Ltd., London, 1912.
- Colburn, D. S., and C. S. Sonett, Discontinuities in the solar wind, Space Sci. Rev., **5**, 439, 1966.
- Csoeke-Poeckh, A., R. H. Lee, W. J. Wagner, L. L. House, E. Hildner, and C. Sawyer, Coronal observations from the SMM satellite, Journal of Spacecraft, **19**, 345, 1982.
- Dryer, M., Solar-generated disturbances and their propagation through the interplanetary medium, Pp. 199 in Solar Wind Four, Max Planck Institute report MPAE-W--100-81-31, 1981.
- Dryer, M., Coronal transient phenomena, Space Sci. Rev., **33**, 233, 1982.
- Dryer, M., and A. Maxwell, Radio data and a theoretical model for the fast-mode MHD shock wave generated by the solar flare of 1973 September 5, 1826 UT, Astrophys. J., **231**, 945, 1979.
- Dryer, M., S. T. Wu, R. S. Steinolfson, and R. M. Wilson, Magnetohydrodynamic models of coronal transients in the meridional plane. II. Simulation of the coronal transient of 1973 August 21, Astrophys. J., **227**, 1059, 1979.
- Dulk, G. A., Radio and white-light observations of coronal transients, p 419 in Radio Physics of the Sun (ed. by M. R. Kundu and T. E. Gergely) 1980.
- Feldman, W. C., J. R. Asbridge, S. J. Bame, and J. T. Gosling, Long-term variations of selected wind properties: IMP 6, 7, and 8 results, J. Geophys. Res., **83**, 2177, 1978.
- Feldman, W. C., J. R. Asbridge, S. J. Bame, E. E. Fenimore, and J. T. Gosling, The solar origins of solar wind interstream flows: Near-equatorial coronal streamers, J. Geophys. Res., **86**, 5408, 1981.
- Fenimore, E. E., Solar wind flows associated with hot heavy ions, Astrophys. J., **235**, 245, 1980.
- Fisher, R. R., and A. I. Poland, Coronal activity below 2 Ro: 1980 February 15-17, Astrophys. J., **246**, 1004, 1981.
- Fisher, R. R., R. H. Lee, R. M. MacQueen, New Mauna Loa coronagraph systems, Applied Optics, **20**, 1094, 1981.

- Gary, D. E., G. A. Dulk, L. L. House, R. M. E. Illing, C. Sawyer, D. J. McLean, and E. Hildner, Type II bursts, shock waves, and coronal transients: the event of 1980 June 29, 0233 UT, preprint, 1982.
- Gergely, T. E., and M. R. Kundu, Decameter type IV bursts associated with coronal transients, Solar Phys., **34**, 433, 1974.
- Gergely, T. E., M. R. Kundu, and E. Hildner, A coronal transient associated with a high-speed type II burst, Astrophys. J., **268**, 403, 1983a.
- Gergely, T. E., M. R. Kundu, F. T. Erskine III, C. Sawyer, W. J. Wagner, R. M. E. Illing, L. L. House, M. K. McCabe, R. T. Stewart, G. J. Nelson, M. J. Koomen, D. Michels, R. Howard, and N. Sheeley, Radio and visible light observations of a coronal arcade transient, to be published in Solar Physics, 1983b.
- Gold, T., p. 103 in Gas Dynamics of Cosmic Clouds (ed. by H. C. Van de Hulst and J. M. Burgers), North-Holland, Amsterdam, 1955.
- Gold, T., p. 162 of the Solar wind (ed. by R. J. Mackin and M. Neugebauer), Pergamon Press, New York, 1966.
- Gosling, J. T., Large-scale inhomogeneities in the solar wind of the solar origin, Rev. Geophys. Space Sci., **13**, 1053, 1975.
- Gosling, J. T., Transient phenomena in the solar atmosphere and solar wind, p 286 in Physics of Solar Planetary Environments (ed. by D. J. Williams), A.G.U., 1976.
- Gosling, J. T., E. Hildner, R. M. MacQueen, R. H. Munro, A. I. Poland, Mass ejections from the sun: A view from Skylab, J. Geophys. Res., **79**, 4581, 1974a.
- Gosling, J. T., V. Pizzo, and S. J. Bame, Anomalous low proton temperatures in the solar wind following interplanetary shock waves; Evidence for magnetic bottles? J. Geophys. Res., **78**, 2001, 1974b.
- Gosling, J. T., E. Hildner, R. M. MacQueen, R. H. Munro, A. I. Poland, and C. L. Ross, Direct observation of a flare-related coronal and solar wind disturbance, Solar Phys., **40**, 439, 1975.
- Gosling, J. T., E. Hildner, R. M. MacQueen, R. H. Munro, A. I. Poland, and C. L. Ross, The speeds of coronal mass ejection events, Solar Phys., **48**, 389, 1976.
- Gosling, J. T., E. Hildner, J. R. Asbridge, S. J. Bame, and W. C. Feldman, Noncompressive density enhancements in the solar wind, J. Geophys. Res., **82**, 5005, 1977.
- Gosling, J. T., J. R. Asbridge, S. J. Bame, W. C. Feldman, and R. D. Zwickl, Observations of large fluxes of  $\text{He}^+$  in the solar wind following an interplanetary shock, J. Geophys. Res., **85**, 3431, 1980.
- Gosling, J. T., G. Borrini, J. R. Asbridge, S. J. Bame, W. C. Feldman, and R. T. Hansen, Coronal streamers in the solar wind at 1 AU, J. Geophys. Res., **86**, 5438, 1981.
- Greaves, W. M. H., and H. W. Newton, On the recurrence of magnetic storms, Mon. Not. Roy. Astron. Soc. London, **89**, 641, 1929.

- Hildner, E., Mass ejections from the corona into interplanetary space, p. 3 of Study of Traveling Interplanetary Phenomena (ed. by M. A. Shea, D. Smart, and S. T. Wu), D. Reidel, Dordrecht, 1977.
- Hildner, E., J. T. Gosling, R. M. MacQueen, R. H. Munro, A. I. Poland, and C. L. Ross, Frequency of coronal transients and solar activity, Solar Phys., **48**, 127, 1976.
- Hirshberg, J., The transport of plasma from the sun to the earth, Planet Space Sci., **16**, 309, 1968.
- Hirshberg, J., and D. S. Colburn, Interplanetary field and geomagnetic variations: a unified view, Planet. Space Sci., **17**, 1183, 1969.
- Hodgson, R., On a curious appearance seen in the sun, Mon. Not. R. Astron. Soc., **20**, 15, 1860.
- House, L. L., W. J. Wagner, E. Hildner, C. Sawyer, and H. U. Schmidt, Studies of the corona with the Solar Maximum Mission Coronagraph/Polarimeter, Astro Phys. J., **244**, L 117, 1981.
- Howard, R. A., material presented at this workshop.
- Howard, R. A., M. J. Koomen, D. J. Michels, R. Tousey, C. R. Detwiler, D. E. Roberts, R. T. Seal, and J. D. Whitney, NOAA World Data Center A for Solar-Terrestrial Physic, Report No. UAG-484, 1976.
- Howard, R. A., D. J. Michels, N. R. Sheeley, and M. J. Koomen, The observation of a coronal transient directed at earth, Astrophys. J., **203**, L 101, 1982.
- Hundhausen, A. J., Coronal Expansion and Solar Wind, Springer-Verlag, Heidelberg, 1972
- Hundhausen, A. J., S. J. Bame, and M. D. Montgomery, The large-scale characteristics of flare-associated solar wind disturbances, J. Geophys. Res., **75**, 4631, 1970.
- Hundhausen, A. J., C. B. Sawyer, L. L. House, R. M. E. Illing, and W. J. Wagner, Coronal mass ejections observed during the Solar Maximum Mission: Latitude distribution and rate of occurrence, submitted to J. Geophys. Res., July, 1983.
- Illing, R. M. E., and A. J. Hundhausen, Possible observation of a disconnected magnetic structure in a coronal mass ejection, submitted to J. Geophys. Res., July, 1983.
- Kahler, S., The morphological and statistical properties of solar X-ray events with long decay times, Astrophys. J., **214**, 891, 1977.
- Klein, L. W., and L. F. Burlaga, Interplanetary magnetic clouds at 1 AU, J. Geophys. Res., **87**, 613, 1982.
- Kreplin, R. W., T. A. Chubb, and H. Friedman, X-ray and Lyman-alpha emission from the sun as measured from the NRL SR-1 satellite, J. Geophys. Res., **65**, 2231, 1962.
- Kundu, M. R., Solar Radio Astronomy, Interscience, New York, 1965.
- Kutchko, F. J., P. R. Briggs, and T. D. Armstrong, The bi-directional particle event of October 12, 1977, possibly associated with a magnetic loop, J. Geophys. Res., **87**, 1419, 1982.

- Low, B. C., Eruptive solar magnetic fields, Astrophys. J., **251**, 352, 1981.
- Low, B. C., Self-similar magnetohydrodynamics I. The  $\gamma = 4/3$  polytrope and the coronal transient, Astrophys. J., **254**, 796, 1982a.
- Low, B. C., Self-similar magnetohydrodynamics II. The expansion of a stellar envelope into a surrounding vacuum, Astrophys. J., **261**, 351, 1982b.
- Low, B. C., R. H. Munro, and R. R. Fisher, The initiation of a coronal transient, Astrophys. J., **254**, 335, 1982.
- MacQueen, R. M., Coronal transients: a summary, Phi. Trans. R., Soc. Lond., **A 297**, 605, 1980.
- MacQueen, R. M., and R. R. Fisher, Kinematics of solar inner coronal transients, submitted to Solar Physics, 1983.
- MacQueen, R. M., J. T. Gosling, E. Hildner, R. H. Munro, A. I. Poland, and C. L. Ross, Proceedings of the Society of Photo Optical Engineers, **44**, 207, 1974.
- Malitson, H. H., J. Fainberg, and R. G. Stone, A density scale for the interplanetary medium from observations of a type II solar radio burst out to 1 astronomical unit, Astrophys. J., **L 35**, 1973.
- Maxwell, A., and M. Dryer, Solar radio bursts of spectral type II, coronal shocks, and optical coronal transients, Solar Phys., **73**, 313, 1981.
- Maxwell, A., and M. Dryer, Characteristics of shocks in the solar corona as inferred from radio, optical, and theoretical investigations, Space Sci. Rev., **32**, 11, 1982.
- Montgomery, M. D., J. R. Asbridge, S. J. Bame, and W. C. Feldman, Solar wind electron temperature depressions following some interplanetary shock waves: evidence for magnetic merging? J. Geophys. Res., **79**, 2324, 1974.
- Moore, R., D. L. McKenzie, Z. Svestka, K. G. Widing, S. K. Antiochos, K. P. Dere, H. W. Dodson-Prince, E. Hiei, K. R. Krall, A. S. Krieger, H. E. Mason, R. D. Petrasso, G. W. Pneuman, J. K. Silk, J. A. Vorpahl, and G. L. Withbroe, The thermal X-ray flare plasma, in Solar Flares, edited by P. A. Sturrock, p. 341, Colorado Associated University Press, Boulder, CO, 1980.
- Mouschovia, T. Ch., and A. I. Poland, Expansion and broadening of coronal loop transients: a theoretical explanation, Astrophys. J., **220**, 675, 1978.
- Munro, R. H., J. T. Gosling, E. Hildner, R. M. MacQueen, A. I. Poland, and C. L. Ross, The association of coronal mass ejection transients with other forms of solar activity, Solar Phys., **61**, 201, 1979.
- Nakagawa, Y., and R. S. Steinolfson, Dynamical response of the solar coronal I. Basic formulations, Astrophys. J., **207**, 296, 1976.
- Nakagawa, Y., S. T. Wu, and E. Tandberg-Hanssen, Dynamic response of an isothermal static coronal to finite-amplitude disturbances, Solar Phys., **41**, 387, 1975.

- Nakagawa, Y., S. T. Wu, and S. M. Han, Magnetohydrodynamics of atmospheric transients, II. Two-dimensional numerical results for a model solar corona, Astrophys. J., **219**, 314, 1978.
- Nakagawa, Y., S. T. Wu, and S. M. Han, Dynamics of coronal transients: Two-dimensional, non-plane, MHD models, p. 495 in Solar and Interplanetary Dynamics, I.A.U., Symp. 91 (ed. by M. Dryer and E. Tandberg-Hanssen), D. Reidel, Dordrecht, 1980.
- Nakagawa, Y., S. T. Wu, and S. M. Han, Magnetohydrodynamics of atmospheric transients, III. Basic results of nonplane two-dimensional analysis, Astrophys. J., **244**, 331, 1981.
- Neugebauer, M., Observational constraints on solar wind acceleration mechanisms, to be published in Solar Wind V (ed. by M. Neugebauer) NASA-SP, 1983.
- Palmer, I. D., F. R. Allum, and S. Singer, Bidirectional anisotropies in solar cosmic ray events: evidence for magnetic bottles, J. Geophys. Res., **83**, 75, 1978.
- Parker, E. N., Sudden expansion of the corona following a large solar flare and the attendant magnetic field and cosmic ray effects, Astrophys. J., **133**, 1014, 1961.
- Parker, E. N., Interplanetary Dynamical Processes, Interscience, New York, 1963.
- Pneuman, G. W., Reconnection driven control transients, p. 317 in Solar and Interplanetary Dynamics, IAU Symp. 91, (ed. by M. Dryer and E. Tandberg-Hanssen), Dordrecht, 1980a.
- Pneuman, G. W., Eruptive prominences and coronal transients, Solar Phys., **65**, 309, 1980b.
- Pneuman, G. W., Diamagnetic aspects of the coronal transient phenomenon, Adv. Space Res., **2**, 233, 1983.
- Pudovkin, M. I., S. A. Zaitseva, and E. E. Berevslenska, The structure and parameters of flare streams, J. Geophys. Res., **84**, 6649, 1979.
- Rosner, R., B. C. Low, and T. E. Holzer, Physical processes in the solar corona, to be published in Physics of the Sun, 1983.
- Rust, D. M., Coronal disturbances and their terrestrial effects, Space Sci. Rev., **34**, 21, 1983.
- Rust, D. M., E. Hildner, M. Dryer, R. T. Hausen, A. N. McClymont, S. M. P. McKenna Lawlor, D. J. McLean, E. J. Schmahl, R. S. Steinolfson, E. Tandberg-Hanssen, R. Tousey, D. F. Webb, and S. T. Wu, Mass ejections, p. 273 in Solar Flares (ed. by P. A. Sturrock), Colorado Associated University Press, Boulder, 1980.
- Sawyer, C., W. J. Wagner, A. J. Hundhausen, L. L. House, and R. M. E. Illing, Activity associated with coronal mass ejections at the maximum and minimum of the solar cycle, submitted 1983.
- Schatten, K. H., and J. E. Schatten, Magnetic field structure in flare-associated solar wind disturbances, J. Geophys. Res., **77**, 4858, 1972.
- Schwenn, R., Direct correlations between coronal transients and interplanetary disturbances, Space Sci. Rev., **34**, 85, 1983.

- Schwenn, R. H., H. Rosenbauer, and K. H. Muhlhauser, Singly-ionized helium in the driver gas of interplanetary shock waves, Geophys. Res. Lett., **7**, 201, 1980.
- Sheeley, N. R., Jr., J. D. Bohlin, G. E. Brueckner, J. D. Purcell, V. E. Scherrer, R. Tousey, J. B. Smith, D. M. Speich, E. Tandberg-Hanssen, R. M. Wilson, A. C. DeLoach, R. B. Hoover, and J. P. McGuire, Coronal changes associated with a disappearing filament, Solar Phys., **45**, 377, 1975.
- Sheeley, N. R., Jr., D. J. Michels, R. A. Howard, and M. J. Koomen, Initial observations with the Solwind coronagraph, Astrophys. J., **237**, L 99, 1980.
- Sheeley, N. R., Jr., R. A. Howard, M. J. Koomen, D. J. Michels, K. L. Harvey, and J. W. Harvey, Observations of coronal structure during sunspot maximum: Space Sci. Rev., **33**, 219, 1982a.
- Sheeley, N. R., Jr., R. A. Howard, M. J. Koomen, D. J. Michels, and A. I. Poland, The observation of a high-latitude coronal transient, preprint, 1982b.
- Sheeley, N. R., Jr., R. A. Howard, M. J. Koomen, and D. J. Michels, Associations between coronal mass ejections and X-ray events, Astrophys. J., in press, 1983a.
- Sheeley, N. R., Jr., R. A. Howard, M. J. Koomen, D. J. Michels, R. Schwenn, K. H. Muhlhauser, and H. Rosenbauer, Associations between coronal mass ejections and interplanetary shocks, to be published in the Proceeding of the 5th Solar Wind Conference, 1983b.
- Steinolfson, R. S., Coronal loop transients in streamer configurations, Astron. Astrophys., **115**, 39, 1982a.
- Steinolfson, R. S., Coronal response to a solar event in a corona evacuated by a prior transient, Astron. Astrophys., **115**, 50, 1982b.
- Steinolfson, R. S., and Y. Nakagawa, Dynamical response of the corona, II. Numerical simulations near the sun, Astrophys. J., **207**, 300, 1976.
- Steinolfson, R. S., and Y. Nakagawa, Dynamical response of the solar corona, III. Numerical simulation of the 1973 June 10 coronal transient, Astrophys. J., **215**, 345, 1977.
- Steinolfson, R. S., S. T. Wu, M. Dryer, and E. Tandberg-Hanssen, The effect of the magnetic field on coronal transients, p. 23 in Proceedings of COSPAR Symposium B, AFGL-TR-77-0309, 1977.
- Steinolfson, R. S., S. T. Wu, M. Dryer, and E. Tandberg-Hanssen, Magnetohydrodynamic models of coronal transients in the meridional plane. I. The effect of the magnetic field, Astrophys. J., **225**, 259, 1978.
- Steinolfson, R. S., E. J. Schmahl, and S. T. Wu, Hydrodynamic simulations of flare/surge events, Solar Phys., **63**, 187, 1979.
- Steinolfson, R. S., S. T. Wu, M. Dryer, and E. Tandberg-Hanssen, Effect of emerging magnetic flux on the solar corona, p. 51 in Solar Wind Four, Max Plank Institut. Tech Report No. MPAE-W-100-81-31, 1981.

- Stewart, R. T., R. A. Howard, F. Hanssen, T. Gergely, and M. Kundu, Observations of coronal disturbances from 1 to 9  $R_{\odot}$ , Solar Phys., **36**, 219, 1974.
- Stewart, R. T., Transient disturbances of the corona, p. 333 in Solar and Interplanetary Dynamics (ed. by M. Dryer and E. Tandberg-Hanssen) Reidel, Dordrecht, 1980.
- Taylor, H. E. Sudden commencement associated discontinuities in the interplanetary magnetic field observed by Imp 3, Solar Phys., **6**, 320, 1969.
- Temny, V. V., and O. L. Vaisberg, Dumb-bell distributions of suprathermal solar wind electrons from Prognoz-7 observations, Space Research Institute II, p 499, preprint, 1979.
- Teske, R. G., Soft X-rays and solar activity, III. Loop prominences with soft X-ray emission, Solar Phys., **17**, 76, 1971.
- Tousey, R., The solar corona, Space Res., **13**, 713, 1973.
- Vaiana, G. S., and R. Rosner, Recent advances in coronal physics, Ann Rev. Astron. Astrophys., **16**, 339, 1978.
- Wagner, W. J., SERF studies of mass motions arising in flares, Adv. Space Res., **2**, 203, 1983.
- Wagner, W. J., R. M. E. Illing, C. B. Sawyer, L. L. House, N. R. Sheeley, R. A. Howard, M. J. Koomen, D. J. Michels, R. N. Smartt, and M. Dryer, A white light/Fe X/H coronal transient observation to 10 solar radii, Solar Phys., **83**, 153, 1983.
- Webb, D. F., A. S. Krieger, and D. M. Rust, Coronal X-ray enhancements associated with H $\alpha$  filament disappearances, Solar Phys., **48**, 159, 1976.
- Wild, J. P., and S. F. Smerd, Radio bursts from the solar corona, Ann. Rev. Astron. Astrophys., **10**, 159, 1972.
- Wild, J. P., S. F. Smerd, and A. A. Weiss, Solar bursts, Ann Rev. Astron. Astrophys., **1**, 291, 1963.
- Woo, R., J. W. Armstrong, N. R. Sheeley, Jr., R. A. Howard, D. J. Michels, M. J. Koomen, Simultaneous radio scattering and white light observations of a coronal transient, Nature, **300**, 157, 1982.
- Wu, S. T., Theoretical interpretation of traveling interplanetary phenomena and their solar origins, p. 443 in Solar and Interplanetary Dynamics, IAU Symp. 91, (ed. by M. Dryer and E. Tandberg-Hanssen), D. Reidel, Dordrecht, 1980.
- Wu, S. T., Numerical simulation of magnetohydrodynamic shock propagation in the corona, Space Science Reviews, **32**, 115, 1982.
- Wu, S. T., M. Dryer, Y. Nakagawa, and S. M. Han, Magnetohydrodynamics of atmospheric transients. II. Two-dimensional numerical results for a model solar corona, Astrophys. J., **219**, 324, 1978.

- Wu, S. T., L. C. Kan, Y. Nakagawa, and E. Tandberg-Hanssen, Effect of thermal conduction and radiation on the dynamics of a flaring coronal loop, Solar Phys., **70**, 137, 1981a.
- Wu, S. T., R. S. Steinolfson, M. Dryer, and E. Tandberg-Hanssen, Magnetohydrodynamic models of coronal transients in the meridional plane. IV. Effects of the solar wind, Astrophys. J., **243**, 641, 1981b.
- Wu, S. T., Y. Nakagawa, S. M. Han, and M. Dryer, Magnetohydrodynamics of atmospheric transient, IV. Nonplane two-dimensional analyses of energy conversion and magnetic field evolution, Astrophys. J., **262**, 369, 1982.
- Wu, S. T., S. Wang, M. Dryer, A. L. Poland, D. G. Sime, C. J. Wolfson, L. E. Orwig, and A. Maxwell, Magnetohydrodynamic simulation of the coronal transient associated with the solar limb flare of 1980 June 29, 1821 UT, to be published in Solar Phys., 1983.
- Yeh, T., A Magnetohydrodynamic theory of coronal loop transients, Solar Physics, **78**, 287, 1982.
- Yeh, T., Diamagnetic force on a flux tub, Astrophys. J., **264**, 630, 1983.
- Yeh, T., and M. Dryer, Effects of self-induced magnetic force in a coronal loop transient, Astrophys. J., **245**, 704, 1981a.
- Yeh, T., and M. Dryer, An acceleration mechanism for loop transients in the outer corona, Solar Phys., **71**, 141, 1981b.
- Zwickl, R. D., J. R. Asbridge, S. J. Bame, W. C. Feldman, and J. T. Gosling, He<sup>+</sup> and other unusual ions in the solar wind: A systematic search covering 1972-1980, J. Geophys. Res., **87**, 7379, 1982.
- Zwickl, R. D., J. R. Asbridge, S. J. Bame, W. C. Feldman, J. T. Gosling, and E. J. Smith, Plasma properties of driver gas following interplanetary shocks observed by ISEE-3, in solar Wind V, edited by M. Neugebauer, NASA SP to be published in 1983.



## CHAPTER 7

# CONNECTION BETWEEN THE MAGNETOSPHERE AND IONOSPHERE

### WORKING GROUP MEMBERS

Y. T. Chiu, Chairman  
*The Aerospace Corporation*

R. Anderson  
*University of Iowa*

J. Fennell  
*The Aerospace Corporation*

L. Frank  
*University of Iowa*

R. Hoffman  
*NASA/Goddard Space Flight Center*

M. Hudson  
*University of California, Berkeley*

L. Lyons  
*NASA/Marshall Space Flight Center*

P. Palmadesso  
*Naval Research Laboratory*

E. Ungstrup  
*Danish Space Research Institute*

R. Vondrak  
*Lockheed Palo Alto Research Laboratory*

D. Williams  
*Johns Hopkins University*

R. Wolf  
*Rice University*

**CHAPTER 7**  
**CONNECTION BETWEEN THE MAGNETOSPHERE AND IONOSPHERE**

I.	Introduction	7-3
II.	Global Scale Coupling	7-4
III.	Intermediate Scale Coupling	7-18
IV.	Microscopic Scale Effects	7-29
V.	Summary/Conclusions	7-34
VI.	References	7-36

## I. INTRODUCTION

Two decades of space research have produced ample evidence that particles and fields originating in the active Sun can gain entry into the terrestrial magnetosphere and deposit their energy in the ionosphere and atmosphere. The final link in this solar-terrestrial chain is generically referred to as magnetosphere-ionosphere coupling (MIC). Included in the physics of the MIC processes are not only plasma processes associated with the earthward deposition of energy of the active Sun but also the response of the ionosphere-atmosphere system to the energy deposition. Because of its key position in the physics of the near-Earth space environment, MIC processes have been a major focus of space research.

Up until the past several years, the important issue in MIC was the role of the ionosphere in magnetospheric field configuration (Axford and Hines, 1961; Dungey, 1961) of the global scale. Through comprehensive formulations of MIC theory (Fejer, 1964; Vasyliunas, 1970; Stern, 1977; Wolf, 1975), the schematic relationships between magnetospheric magnetic and electric configurations and ionospheric plasma flow have been elucidated. Thus, the ionosphere-atmosphere, through dissipative processes inherent in the finite ionospheric conductivity, is thought to regulate the electrodynamic circuit driven by the interaction between the magnetosphere and the solar wind at the magnetospheric boundary.

Despite the highly successful elucidation of the schematic features of MIC by global-scale theory, it soon became clear that MIC involves much more complex plasma kinetic effects. Evidences for this significant change of theoretical perspective were made available since the latter part of the 1970s by plasma and field observations of auroral satellites in the  $\sim 1 R_E$  altitude region and by rocket chemical releases. New observations from these auroral satellites have not only confirmed previous suspicions that a parallel electric potential drop may interface with global MIC processes but have also introduced entirely new kinetic phenomena such as up-flowing ionospheric ions and complex wave-particle interactions, which are likely to change our perspective of the role of the topside ionosphere in MIC processes. Recent observations of polar region auroral phenomena have indicated that our picture of magnetospheric topology may need to be revised. In short, we have witnessed in the last 4-5 years a dramatic change in perspective in the physics of magnetosphere-ionosphere coupling.

Because of the far-reaching implications of recent discoveries in MIC and because a coherent assessment of them has yet to be made, the principal task of our group is to critically assess these new observations and the new perspectives that they may engender. Thus, we shall not devote much effort to reviewing the general scheme of MIC prior to the latter part of the 1970s. The reader can be referred to a number of excellent reviews on the subject (e.g., Banks, 1979 and references therein).

Scientific advances cannot easily be categorized. This is especially true in a cross-disciplinary subject undergoing dynamic change; therefore, in our consideration of new MIC processes, we shall not strive for a consensus view on any topic; rather, we shall strive for a balanced presentation of viable alternate perspectives. Our multilateral considerations do not, however, imply that the new discoveries in MIC processes are a haphazard collection of facts open to a host of interpretations. Indeed, we shall demonstrate that the new kinetic phenomena are infused with unifying themes which correlate plasma processes of vastly different scales in time and space into a multiplex of couplings, rather than into a monolithic coupling process. Thus, the observed multiplicity of space-time scales, of new magnetospheric configurations and of new plasma sources is itself a major new perspective in space physics. In the following, considerations of these important new topics are given as if in isolation from each other, but the reader is reminded to pay special attention to the essential links between topics.

In this report, MIC topics of interest are roughly grouped according to the scale lengths of the phenomena being treated. This particular choice of groupings is mainly for convenience, with perhaps some suggestion as to the direction of energy cascade from the largest scales down to the smaller scales.

## II. GLOBAL SCALE COUPLING

In this section, we consider plasma and field phenomena relevant to structures of the MIC process on the global scale. Some of these processes may, however, have their origin in microscopic kinetic plasma interactions.

### MAGNETOSPHERIC CONFIGURATION EFFECTS

The first two panels in Figure 7-1 represent the basic configuration of electrical coupling between the magnetosphere and the ionosphere. The first panel is a cross-sectional view of the magnetospheric magnetic and electric fields in the dawn-dusk meridional plane. This magnetospheric electric field is mapped along the magnetic field lines onto the terrestrial ionosphere, which results in the  $\vec{E} \times \vec{B}$  convection pattern of ionospheric plasma shown in panel B. The convection pattern normally has two cells with antisunward flow in the polar cap and sunward flow at lower latitudes near the auroral oval. It has recently been determined that considerable distortions of this pattern can occur, and more complicated multicelled patterns have been inferred from measurements. Furthermore, the coupling between the magnetosphere and the ionosphere is affected by the presence of parallel electric fields in auroral arcs and in the polar cap.

NOMINAL MAGNETOSPHERE-IONOSPHERE COUPLING

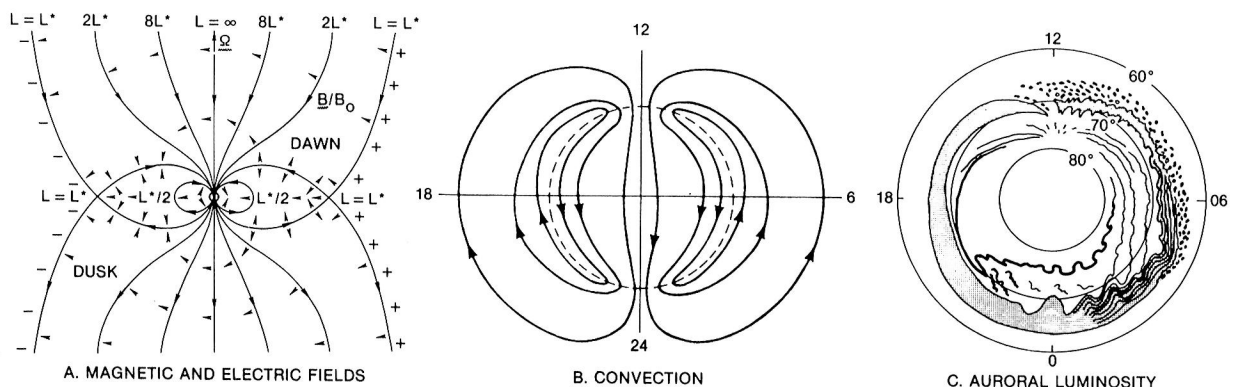


Figure 7-1: Nominal features of magnetospheric-ionospheric coupling. (a) A schematic illustration of magnetospheric structure showing closed and open field lines and their separatrices  $L^*$  (Chiu et al., 1983b). The directions of the magnetospheric electric field are shown by arrowheads attached to field lines. (b) High-latitude ionospheric plasma convection flow showing the configuration of the boundary  $L^*$  (dashed line) (Heelis and Hanson, 1980). (c) Auroral morphology (courtesy Akasofu and Kan, 1981).

The electrodynamic coupling between the magnetosphere and ionosphere results in energy dissipation in the ionosphere. Intense Joule heating of the ionosphere occurs in the pre-noon and pre-midnight sectors where the vector difference between the convecting ions and the moving neutral upper atmospheric particles is maximum, resulting in momentum exchange from collisions. The other important process of magnetospheric energy deposition into the upper atmosphere is by precipitation of electrons. The third panel of Figure 7-1 schematically represents the spatial distribution of this energy deposition, as manifested by the distribution of auroral luminosity (Akasofu, 1976). There is a broad belt of electron precipitation that typically occurs in an oval region about the geomagnetic poles at geomagnetic latitudes between  $60^\circ$  and  $80^\circ$ . In the pre-midnight sector bright narrow arcs occur poleward of the widespread diffuse aurora. In the morning sector the precipitation is more structured with the possibility of Sun-aligned arcs at the poleward edge of the auroral oval. The morphology in this figure was deduced primarily from ground-based all-sky camera photographs and generally indicates the regions of electron precipitation from the plasma sheet. Other manifestations of energy input by magnetosphere-ionosphere coupling are now known to be present, including such phenomena as the dayside cusp, polar rain, pulsating aurora, detached arcs in the subauroral evening sector, and stable auroral red (SAR) arcs near the plasmapause.

Recently, global auroral images acquired by the auroral imaging instrument on the Dynamics Explorer-1 spacecraft at high altitudes over the northern polar region have provided details of an auroral configuration which differs significantly from the well-known auroral oval (Frank et al., 1982). The entire auroral oval was at times observed to be divided by a persistent Sun-aligned polar cap arc that intersected the dayside and nightside oval, thus giving the appearance of the Greek letter theta. While Sun-aligned arcs over the Earth's polar caps have been previously reported based on ground-based measurements (Mawson, 1925; Davis, 1962) and from imaging with low-altitude satellites at visible wavelengths (Meng and Akasofu, 1976; Bunn and Shepherd, 1979; Ismail and Meng, 1982), these new observations are unique in their ability to clearly specify the complete connection of the dayside and nightside portions of the oval by the arc and the temporal history of the arc. As viewed at large distances above the northern pole, the polar arc luminosities are not isolated from the auroral oval, but often display brightenings at the intersections of the midday and midnight sectors. The polar arc is observed often to depart from Sun-alignment by as much as  $40^\circ$  and to move across the polar cap with a speed in the range of 100 to 1000 m/s. To date these auroras are found to occur during periods of northward-directed ( $B_z > 0$ ) interplanetary magnetic field.

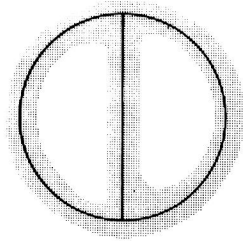
Measurements from the two Dynamics Explorer spacecraft have provided considerable information about the convection and plasma conditions associated with the arc. A four-cell state of the ionospheric convection pattern has been inferred in association with the theta aurora, where a polar sunward convection zone coincides with the position of the polar arc (N. Maynard, private communication). Field-aligned currents are located above the polar arc (M. Sugiura, private communication). Inverted-V electron precipitation events also occur above the polar arc (D. Winningham, private communication) and the plasmas are found to be similar to those associated with the high-latitude portion of the auroral oval (J. Burch and E. Shelley, private communications). Major observational features of the theta auroral are summarized pictorially in the left side of Figure 7.2.

At this time the interpretation of the observations both as to the uniqueness of the form and their implications on magnetospheric topology is under discussion. Previously, Ismail and Meng (1982) used seven years of DMSP imagery from low altitudes to produce a classification system for polar cap auroral arcs which includes polar arcs morphologically connected to the auroral oval at least on the nightside. These arcs, labeled "hook" arcs were also found to occur during periods of northward interplanetary  $B_z$ , with an occurrence frequency of less than one percent. Other Sun-aligned arcs have been interpreted to delineate the poleward expansion of the bright

POLAR IONOSPHERE:

C-682-915

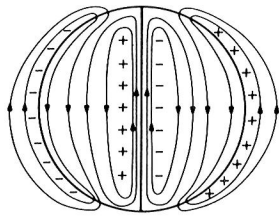
LUMINOSITIES



SUN

INTERPRETIVE FEATURES OF  $\theta$ -AURORAS

CONVECTION



CROSS-SECTION OF MAGNETOTAIL:

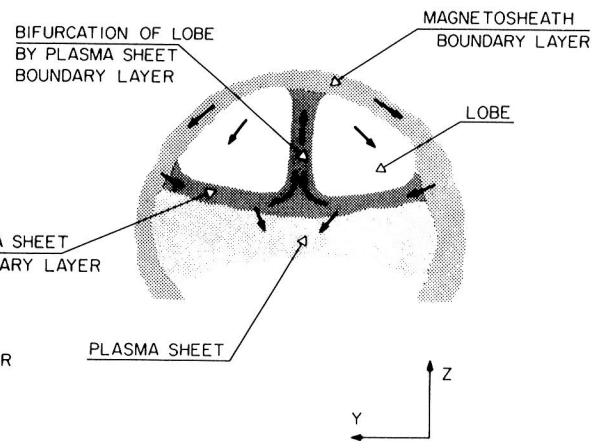
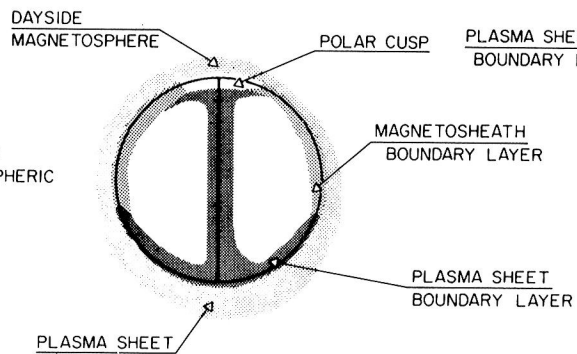
ASSOCIATED  
MAGNETOSPHERIC  
REGIONS

Figure 7-2: Observational and interpretative features of theta auroras.

edge of the evening or morning auroral oval into the center of the polar cap (Ismail and Meng, 1982). A four-cell state of the convection pattern has also been previously related to the northward interplanetary  $B_z$  condition (Burke et al., 1982; Horwitz and Akasofu, 1979). Because the Dynamics Explorer observations reveal the global nature of the arcs, their temporal history, and their striking intensity contrast with the rest of the polar cap, it is not clear whether these earlier observations provided a recognized form of this interesting auroral feature.

The extension of the features of the theta aurora into the magnetosphere is difficult and uncertain. The plasmas over the polar arc are similar to those of the plasma sheet boundary layer that surrounds the plasma sheet in the magnetotail (DeCoster and Frank, 1979). This boundary layer is thought to be topologically connected to the high-latitude nighttime sector of the typical auroral oval. If this association of plasma regimes is correct, the direct implication is that either the magnetotail lobe is bifurcated with a plasma zone to provide for the polar cap arc (right side of Figure 7-2), or a large portion of one side of the lobe is filled with significant plasma with dominant precipitation at the edge of the field lobe to provide the energy for the arc. Thus, an important step in determining the global magnetospheric topology is the search for plasma streams directed northward from the plasma sheet with the ISEE and IMP spacecrafts in the northern magnetotail lobe. Regardless of the accuracy of these preliminary interpretations, it is clear that there is a major reconfiguration of the magnetosphere that appears during the occurrence of theta auroras. Currently there are no mechanisms for bifurcating a magnetotail lobe which can be included in a magnetospheric topological model.

## ELECTRICAL COUPLING

The basic large-scale patterns for ionospheric electric fields, ionospheric currents, and magnetic-field-aligned (Birkeland) currents are now well established (Potemra, 1983).

Polar-orbiting spacecraft and incoherent-backscatter radars have made extensive measurements of high-latitude electric fields. As illustrated in Figure 7-3b, there usually is a dawn-to-dusk electric field across the polar cap, a poleward electric field in the duskside auroral zone, and an equatorward electric field in the dawnside auroral zone. This corresponds to  $\mathbf{E} \times \mathbf{B}$  drift antisunward over the polar caps and sunward in the auroral zones (Figure 7-3d). The pattern strengthens and moves equatorward during periods of high magnetic activity, or when the interplanetary magnetic field is southward. When the IMF is northward, the pattern is weaker but more complicated, often involving a region of sunward flow in the polar cap.

Our observational information about electric currents comes from measurements of magnetic perturbations at the ground and on rockets and satellites. These data have been extensively analyzed to reveal patterns of horizontal ionospheric currents (principally from ground magnetometers) and field-aligned currents (from satellites). With regard to the field-aligned currents, large-scale statistical patterns (locations, intensities, and flow directions) have been determined. The connection of the field-aligned current characteristics with IMF orientation, AE indices, substorm phases, and convective electric fields are also generally known (Potemra, 1979; Saflekos et al., 1982).

The average patterns of ionospheric electric field, flow velocity, Birkeland current, and ionospheric current form a consistent picture, as indicated in Figure 7-3. Meridional Pedersen currents connect region-1 and region-2 currents on the dawn and dusk sides. A westward electrojet of Hall current flows from the dawn side past local midnight, especially during substorms. An eastward electrojet flows on the dusk side. This simple, consistent pattern applies mainly to the large-scale and statistical features (time scales limited by the 100 minute satellite orbital periods).

This simple configuration is complicated by the occurrence of dynamic phenomena and by the presence of small-scale features. Some characteristic features of dynamic events, such as the nighttime westward auroral electrojet associated with substorms, have been identified. Regions where the current configuration and its relations to the electric field, visible aurora, and field-aligned currents have not been fully delineated are the dayside cusp, auroral arcs, the nighttime Harang discontinuity, and the boundary between the auroral oval and the polar cap.

On the theoretical side, much modeling has been done successfully on the ionospheric end of the magnetosphere-ionosphere current system (e.g., Nisbet et al., 1978; Kamide and Matsushita, 1979). They solve the equation of ionospheric-current conservation, given models for the distributions of ionospheric conductivity and Birkeland current. Algorithms have also been developed for deducing the global patterns of ionospheric electric fields and currents and Birkeland currents, given a conductivity model and data from a global array of ground magnetometers (e.g., Kamide et al., 1981).

Turning now to theories that treat magnetospheric and ionospheric currents self-consistently, we note that the patterns shown in Figure 7-3 correspond to the theory of magnetospheric convection, as originally proposed in qualitative form by Axford and Hines (1961) and Dungey (1961). This highly successful theory has still not become quantitative in complete form. There is no full theory of polar-cap convection, due partly to our ignorance about the physics of the far magnetotail. Moving down in latitude, there have been substantial efforts at theories that include boundary layers and their interaction with the ionosphere (e.g., Sonnerup, 1980). However, their accuracy is limited by our ignorance of the detailed physics of the magnetopause. Similarly, numerous theories have coupled the outer plasma sheet to the ionosphere, attempting to illuminate the physics of substorms. Unfortunately, none of these models is widely accepted.

For the inner plasma sheet and the region earthward of it, magnetic field configurations can reasonably be estimated. For this region, Vasyliunas (1970) suggested a computational scheme for treating the ionospheric and magnetospheric ends of a field line self-consistently. This scheme has been developed both in analytic calculations (e.g., Siscoe, 1982) and in a computer model which has been compared with a wide variety of observations (see Wolf et al., 1982 and references therein). The computer model gives reasonably realistic electric fields and currents in the auroral and subauroral regions; it also treats major aspects of the process of ring-current injection. This type of self-consistent model, in which the magnetospheric particles are assumed to undergo gradient, curvature and  $\mathbf{E} \times \mathbf{B}$  drift, and partial rings of magnetospheric current are completed by means of Birkeland currents and ionospheric conduction currents, automatically implies the existence of region-2 Birkeland currents connected to the inner edge of the plasma sheet (first suggested qualitatively by Schield et al., 1969). There is generally good agreement with observations (see Figure 7-4). It should be remarked, however, that there is another line of theoretical argument which neglects gradient/curvature-drift currents, and attributes region-2 currents to inertial effects of plasma slowing down as it moves in from the tail (Rostoker and Bostrom, 1976; Hasegawa and Sato, 1979).

Although there is a fair degree of understanding of large-scale electrical coupling between magnetosphere and ionosphere, many points remain to be clarified. Average patterns are much better understood than dynamic effects and we need more detailed data-theory comparisons for specific times in events. Dynamic, time-dependent effects are important in establishing the relationships between the regions of field-aligned current, convection electric field, and ionospheric conductivity. During disturbed conditions the field-aligned currents and convection electric field penetrate to subauroral latitudes (Klumpar, 1979a; Maynard et al., 1980; Vondrak and Rich, 1982) and effects of the magnetospheric convection field are even observed at the equator (Gonzales et al., 1979).



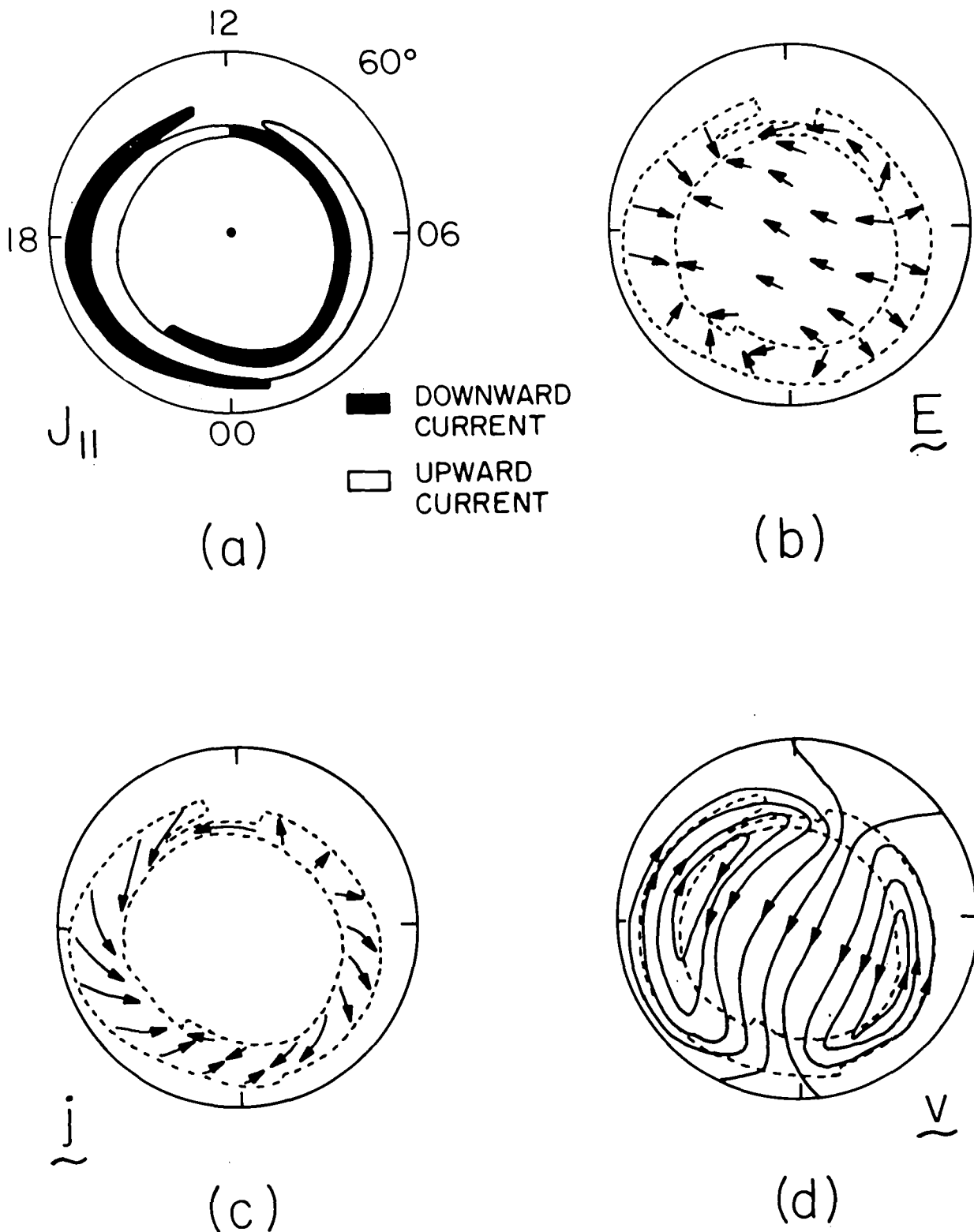


Figure 7-3: Sketch of the typical patterns of (a) Birkeland current (Iijima and Potemra, 1978), (b) electric fields, (c) horizontal ionospheric current, and (d)  $\vec{E} \times \vec{B}$ -drift velocity. Polar cap currents are not shown.

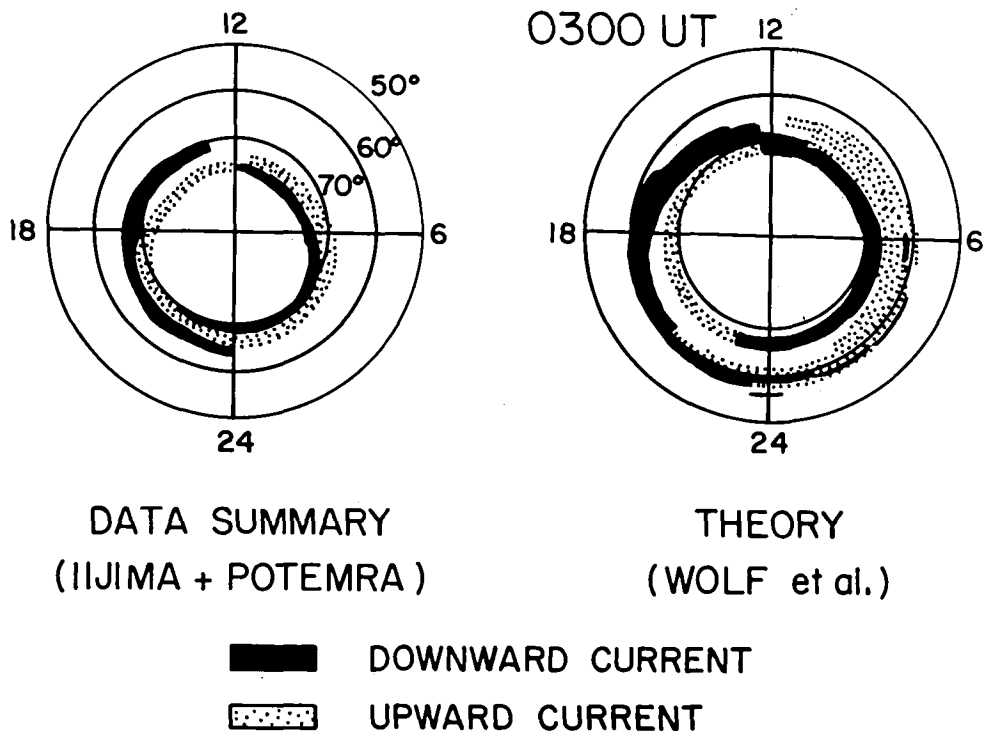


Figure 7-4: Comparison of an observed Birkeland-current pattern (averaged over active times) with a theoretical pattern computed for a specific time during the magnetic storm of July 29, 1977 (Wolf et al., 1982).

## PLASMA TRANSPORT

### Solar Wind and Plasma Sheet

Consideration of magnetosphere-ionosphere coupling processes must include the entry and energization of solar wind plasma in the magnetosphere and its subsequent deposition into the atmosphere. Unlike global scale magnetic and electric field interactions, which primarily control the global configurations of the magnetosphere, the global processes of solar wind plasma interactions in the magnetosphere can directly control and respond to local effects through kinetic processes involving the distribution functions. This is an important theme which runs through a number of phenomena in our discussion.

The magnetosphere has two major plasma sources, the ionosphere and the solar wind (SW). How these sources actually supply the plasma to the magnetosphere and in what proportion are basic problems underlying magnetospheric physics. In MIC we shall only be dealing with the near-Earth part of the problem.

On the dayside solar wind particles enter the magnetosheath and boundary layer (see Haerendel and Paschmann, 1982). Their intrusion into the magnetosphere is not very far and they are intermixed there with plasma of ionospheric origin, especially near the cusp region (Lundin et al., 1982). The SW plasma also has access to the magnetosphere via the distant magnetosheath along the flanks of the magnetosphere and to the distant plasma sheet. The mantle plasma (primarily solar wind mixed with ionospheric ions) gains access to the plasma sheet by convection. The

composition of the plasma sheet varies in response to solar wind conditions and to the energization of the solar wind ions inside the magnetosphere. Solar wind plasma access to the plasma sheet is consistent with a combination of electrostatic and adiabatic processes (Peterson et al., 1981).

An important source of auroral ion precipitation is the acceleration of current-sheet particles by the dawn-to-dusk electric field within the geomagnetic tail. This acceleration process has been investigated by many authors (e.g., Speiser, 1965; Swift, 1977; Cowley, 1980). Lyons and Speiser (1982) calculated the distribution of accelerated ions expected to be ejected from the tail current sheet. They found the ejected ions should form an earthward streaming distribution including all pitch angles, and the pitch angle distribution within the loss cone was calculated to be nearly isotropic at energies of  $\sim 1$  keV to  $>1$  MeV. Satellite measurements of auroral ion precipitation have shown such isotropy at energies of 1-10 keV (Bernstein et al., 1974; Sharber, 1981) and of 115-360 keV (Lundblad et al., 1979). The range of fluxes of precipitating ions reported in the above references are compared in Figure 7-5 with the energy spectra calculated to result from the current sheet acceleration and from observations of earthward streaming ions at the outer boundary of the plasma sheet.

Much of our recent knowledge about inward plasma transport has been obtained by new measurements which include plasma composition and particle distribution functions in two and three dimensions. These have allowed us to see features that were previously hidden. These transported plasmas form the near-Earth plasma sheet and are a source for the ring current. They act on the ionosphere-magnetosphere system via their precipitation into the atmosphere.

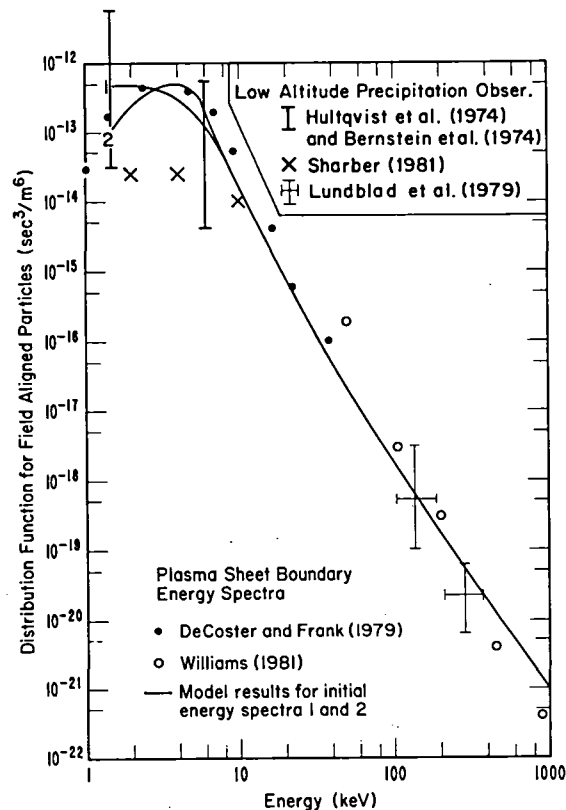


Figure 7-5: Plasma sheet boundary energy spectra. Shown are observations and the model results for the current sheet acceleration using 2 initial distribution functions, models 1 and 2, for the tail mantle plasma. Low-altitude satellite observations of precipitating ions in the auroral zone are also shown.

The convection electric field combined with the geomagnetic field organizes the motions of the particles in these plasmas. Theoretically the particle drift trajectories are organized according to their energy, magnetic moment, pitch angles and species (Cowley and Ashour-Abdalla, 1976a, b; Ejiri, 1978; Kivelson and Southwood, 1975). This has been verified by comparing the observer particle distributions to calculations (Ejiri et al., 1980; Fennel et al., 1981; Y. Chiu, private communications, 1983). Such a comparison of observed and calculated distribution functions is shown in Figure 7-6. The calculations assumed an isotropic plasma sheet distribution at  $12 R_E$  (Y. Chiu, private communications, 1983) and were able to reproduce the dominant features of the observed particle distributions, especially the regions of strong gradients in  $F(v)$ . It has recently been speculated that these convected particle distributions may be unstable. For example, the electron distribution of Figure 7-6 may be unstable to  $(n+1/2)$  electron cyclotron waves (Koons and Fennel, 1983). Loss cone distributions have, for a long time, been promoted as another source of electron cyclotron harmonics (see Kennel and Ashour-Abdalla, 1982). Some observed distributions have been shown to be unstable to  $(n+1/2)f_c$  and upper hybrid emissions (Ronnmark et al., 1978). Recently the observations of a positively sloped feature in the perpendicular distribution function has been identified as a source of free energy for electrostatic waves (Kurth et al., 1980). Similarly, electron and ion distributions on the morningside of the magnetosphere may be unstable to the mirror instability. Such an instability would play a role in morningside aurora formation (Chiu et al., 1983b).

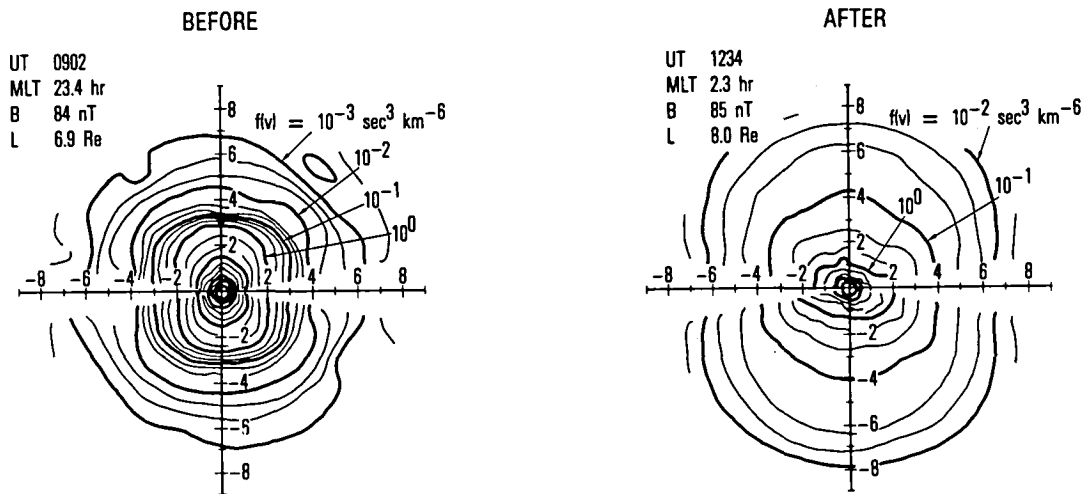
Recently it has been found that energetic (few hundred eV) field-aligned electron beams are observed on the boundary layer of the plasma sheet in the near synchronous region following a plasma injection. These beams are bidirectional (see Figure 7-7) and have been observed on auroral field lines at  $\sim 1 R_E$  altitude (Sharp et al. 1980; Fennel et al., 1981) and deep in the magnetotail (W. Feldman, private communications, 1983). In the near synchronous altitude observations they are also associated with the electrostatic electron cyclotron harmonic (ECH) emissions. ECH emissions increase in intensity and shift from  $nf_c$  to a higher frequency near  $(n+1/2)f_c$ . A pre-existing loss cone distribution appears to be required to generate the ECH waves at  $nf_c$  (B. Edgar, private communication, 1983). An example of this is shown in Figure 7-7 and is to be compared with the expected distribution after injection (Figure 7-6).

Kennel et al. (1970) reported OGO-5 observations of intense electrostatic waves on auroral field lines with the dominant emissions being at a frequency near 1.5 times the electron gyro-frequency. They suggested that these waves may be responsible for the pitch-angle diffusion of auroral region electrons. Using wave vectors predicted theoretically by Young et al. (1973), it was found that the minimum energy for cyclotron resonance with these waves is the order of 0.1 to 1 keV as required to explain diffuse electron precipitation.

A quite general feature of pitch angle diffusion driven by resonant wave-particle interactions is that the diffusion is most intense for the lowest energy particles in resonance with the waves. Resonance extends to higher energies, but the strength of the diffusion decreases with increasing particle energy. Thus particle precipitation into the ionosphere driven by strong, resonant wave-particle interactions should be isotropic in pitch angle at the lowest energies for which precipitation occurs. At increasing energies there should be a transition to anisotropic pitch angle distributions with the anisotropy increasing with increasing particle energy. The energy spectra of the precipitation will show a monotonic decrease with increasing particle energy unless the source magnetospheric particle distribution is peaked in energy.

Lyons (1974) calculated pitch-angle diffusion coefficients for electrons resonant with these electrostatic waves. The results showed that electrons of energies up to a few keV should be put on strong pitch-angle diffusion throughout the loss cone by waves of typically observed amplitudes. Figure 7-8 shows sounding-rocket observations of electron energy spectra at two local pitch angles within the loss cone ( $76^\circ$  and  $42^\circ$ ) obtained during a post-break-up aurora DR

(a) SCATHA/18 APRIL 1979



(b) KINETIC MODEL OF INJECTION

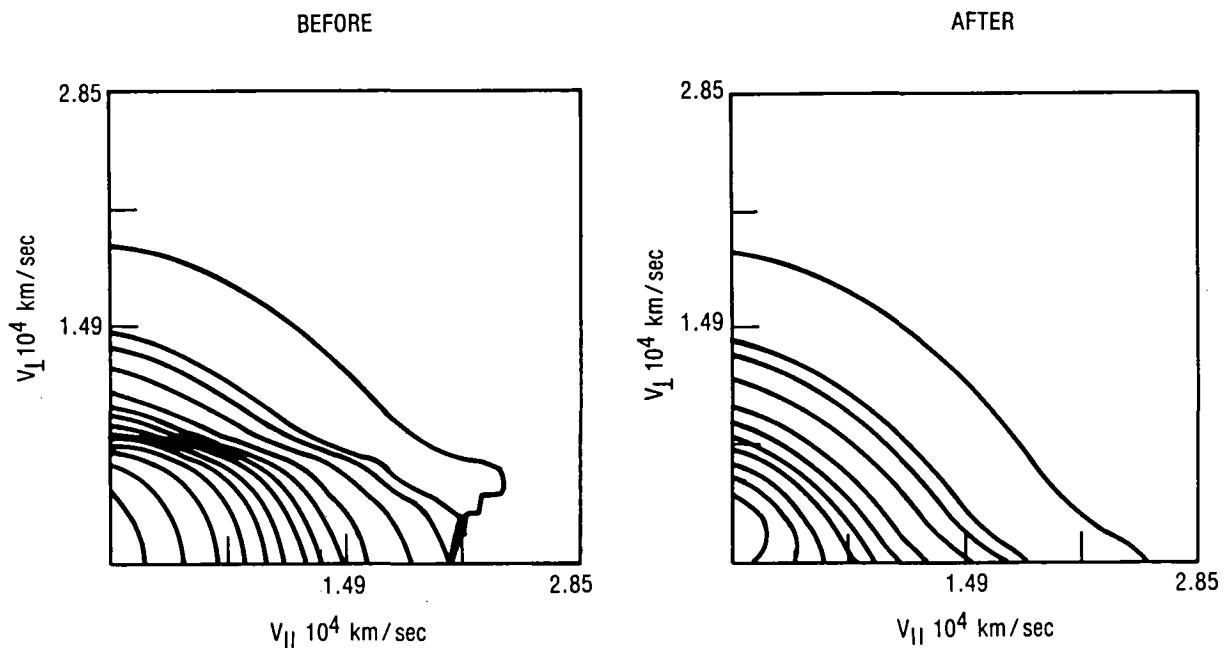


Figure 7-6: (a) Electron distribution signatures near an Alfvén boundary before and after an injection event (Koons and Fennell, 1983). (b) Solutions of Boltzmann's equation before and after a sudden increase in the cross-tail electric field (Chiu and Kishi, private communication, 1983). Note the corresponding features: "box-like" gradients before and field-alignment after the injection.

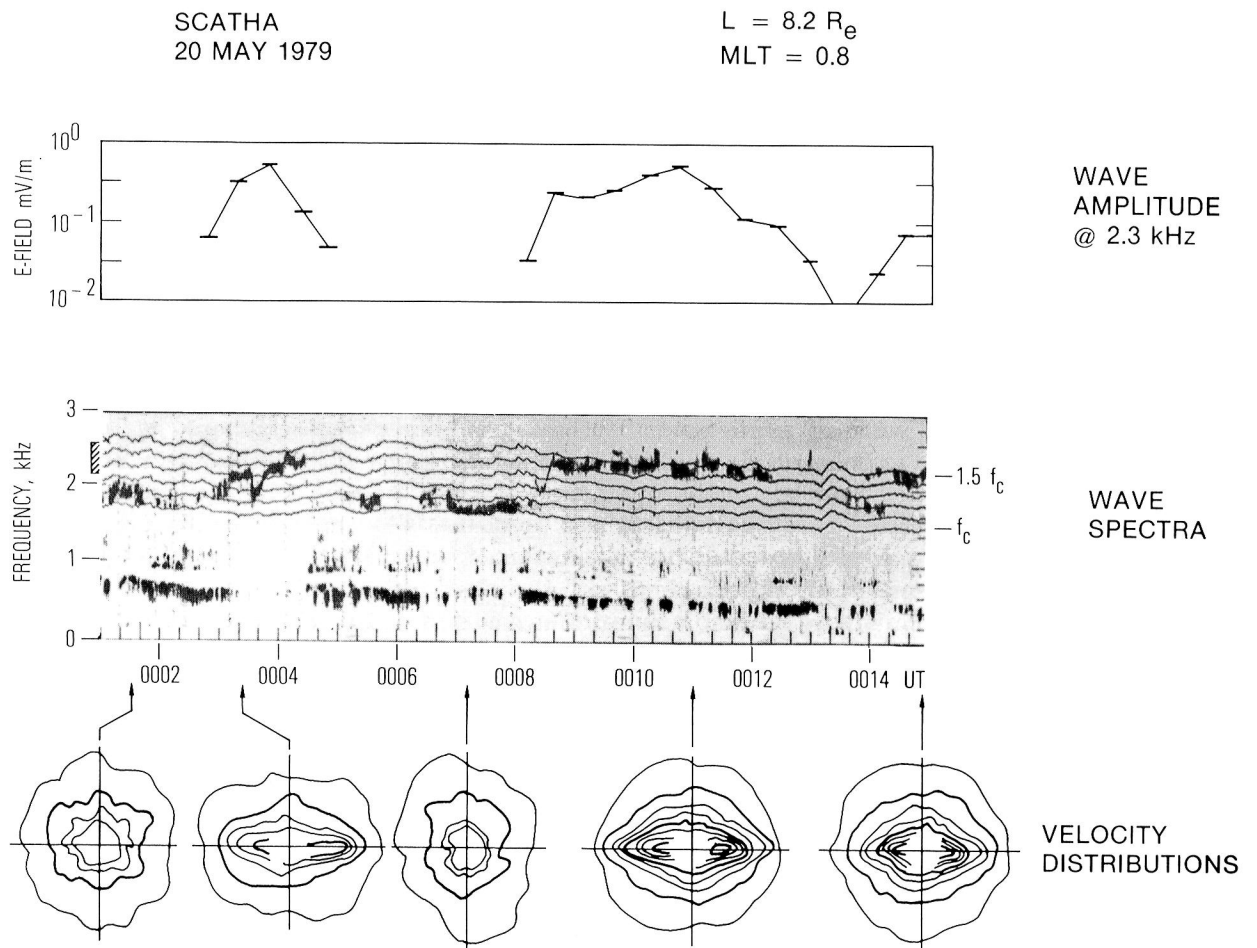


Figure 7-7: Electron cyclotron harmonic emissions (top panel) showing the change in the frequency from  $f_c$  to  $\sim 3/2 f_c$  when the electron distribution changed from a loss cone (bottom left) to a bidirectional field-aligned beam (bottom right) distribution.

(from Whalen and McDiarmid, 1973). Using the observed  $76^\circ$  energy spectra and the calculated diffusion coefficients, Lyons (1974) calculated energy spectra for the  $42^\circ$  pitch-angle particles given by the dashed lines in Figure 7-8.

### Ionospheric Sources

In the past decade measurements from satellites have shown the ionosphere to be an important source of both cold (eV) and hot (keV) ions everywhere in the magnetosphere out to the magnetopause (see review by Horwitz, 1982). Recently new measurements have shown that the ionosphere is an even more extensive source of plasma for the magnetosphere than previously expected. A new warm ( $\sim 10$ -100 eV) component of ionospheric out-flow has been discovered which is even more widespread than the hot source, which is confined to the auroral zone (Shelley et al., 1982). In addition, the polar wind has now been observed at several Earth radii altitude and nitrogen ions have been discovered in the magnetospheric plasma (Gurgiolo and Burch, 1982; Chappell et al., 1982a, b).

In the auroral zone the ionosphere can provide hot plasma ions to the magnetosphere through two mechanisms, both of which involve accelerating ions at relatively low,  $< R_E$ , altitudes. One mechanism energizes the ions perpendicular to the magnetic field (see Figure 7-9b) at relatively low altitudes, generally less than a few thousand kilometers (Shelley et al., 1976; Sharp et al., 1977; Ghielmetti et al., 1978; Gorney et al., 1981; Whalen et al., 1978). There has also been recent evidence for such acceleration higher on the field lines (Horwitz et al., 1982). This source of hot ions has been observed to be of global scale in the auroral region, not limiting itself to any local time sector of the auroral oval (Gorney et al., 1981). More detailed discussions of the mechanism for this new plasma source can be found in Section IV.

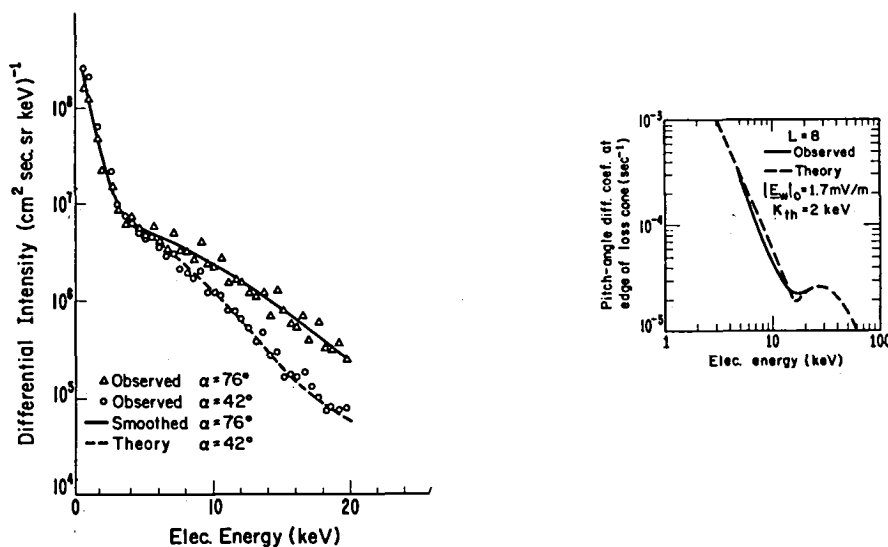


Figure 7-8: Sounding-rocket observations of electron energy spectra obtained within the loss cone at two pitch angles ( $\alpha=76^\circ$  and  $\alpha=42^\circ$ ) during a post-breadup aurora are shown (taken from Whalen and McDiarmid, 1973). In the upper part of the figure, the pitch-angle diffusion coefficient at the edge of the loss cone is shown versus electron energy as obtained from the anisotropy of the observations (solid line) and as obtained from the theoretical calculations for a wave amplitude  $|E_w|_0=1.7$  mV/m and a thermal energy  $K_{th}=2$  keV (dashed line). Also shown, in the main part of the figure is a smooth line through the  $\alpha=76^\circ$  energy spectrum at  $\alpha=42^\circ$  (dashed line) obtained by using the theoretical diffusion coefficients and the smoothed  $\alpha=76^\circ$  spectrum (from Lyons, 1974).

The second mechanism is the field-aligned acceleration of ions by parallel electric fields resulting in an ion beam (Figure 7-9a) with subsequent scattering and thermalization of the flowing plasma (see review by Chiu et al., 1983a and references therein). The combined processes of transverse acceleration followed by field-aligned acceleration at higher altitudes can also occur. This provides in a natural way a hotter ion distribution than would be expected for only field-aligned acceleration of a cold ionospheric plasma. In fact, the higher average energies observed for upward flowing  $O^+$  compared to  $H^+$  (Collin et al., 1981; Lundin et al., 1982) may be a result of preferential transverse acceleration of  $O^+$  compared to  $H^+$  (Ashour-Abdalla et al., 1981).

The observations show that the half-widths of the field-aligned ion (FAI) angular distributions within and above the acceleration region are often of the order of  $10^\circ$ - $45^\circ$  (Richardson et al., 1981; Fennell et al., 1979; Lundin et al., 1982). Given the adiabatic motion of the ions in the geomagnetic field one would predict a strongly collimated distribution at small B values (closer to the equator) with widths on the order of  $5^\circ$  or less. Thus, these ions with broad angular distributions have been "scattered" onto trapped trajectories and have become a permanent component of the plasma sheet. Recently Kaufmann and Kintner (1982) have described a scattering mechanism which would convert parallel ion beam energy to perpendicular energy and exist over a considerable altitude range. In this manner, an ionospheric source localized at the discrete arc can become global in scale.

Examination of the composition of the equatorial FAIs have indicated that they are sometimes dominated by oxygen (see for example Johnson, 1979; Ghielmetti et al., 1979). This is especially true for the FAI component below ring current energies (Kaye et al., 1981). Geiss et al. (1978) estimated that the plasma near synchronous orbit is, on the average, about 50% of terrestrial origin. Peterson et al. (1981) studied the distant plasma sheet and found the terrestrial component to vary from 10% to ~50%. So it appears the plasma sheet has similar composition everywhere and that the ionosphere is a significant source. This source is, however, responding to solar cycle effects (Young et al., 1981).

In addition to contributing to the ion population of the plasma sheet, the ionosphere may contribute significantly to the magnetospheric electron population. Evans and Moore (1979) proposed that plasma sheet electron precipitation in the diffuse aurora produces backscattered and secondary electrons that flow upwards from the ionosphere. While in transit along the magnetic field lines they were pitch-angle scattered onto stably trapped orbits by the same turbulent wave fields responsible for the precipitation. By modeling rocket and satellite measurements of electron energy spectra, they concluded that about 70% of the total electron population of the plasma sheet was of ionospheric origin.

The ionosphere outflow is the mechanism by which the plasmasphere is formed (e.g., Banks, 1979) and by which the polar wind flow is generated (e.g., Banks and Holzer, 1969). In the plasmasphere the resultant high density region is limited in extent because the corotational and solar-wind-induced magnetospheric electric fields cause the plasma to drift on trajectories which do not close about the Earth at larger radial distances than  $5-7 R_E$  at the equator (see Ejiri, 1978 and references therein). At present there are few observations of the thermal ions which make up the polar wind and the plasmasphere (see Horwitz, 1982; Gurgiolo and Burch, 1982). The exact composition and its variations are not well known. It is generally dominated by  $H^+$  with  $He^+$  a minor constituent and very small concentrations of  $He^{++}$ ,  $O^+$ ,  $O^{++}$ ,  $N^+$  and  $N^{++}$  (see review in Horwitz, 1982; Chappell et al., 1982b). Recent observations from chemical releases (Heppner et al., 1981) and photoelectron distributions (Winningham and Gurgiolo, 1982) indicate widespread parallel electric fields over the polar cap. If so, the polar wind outflow will be highly modified (Moore, 1980).



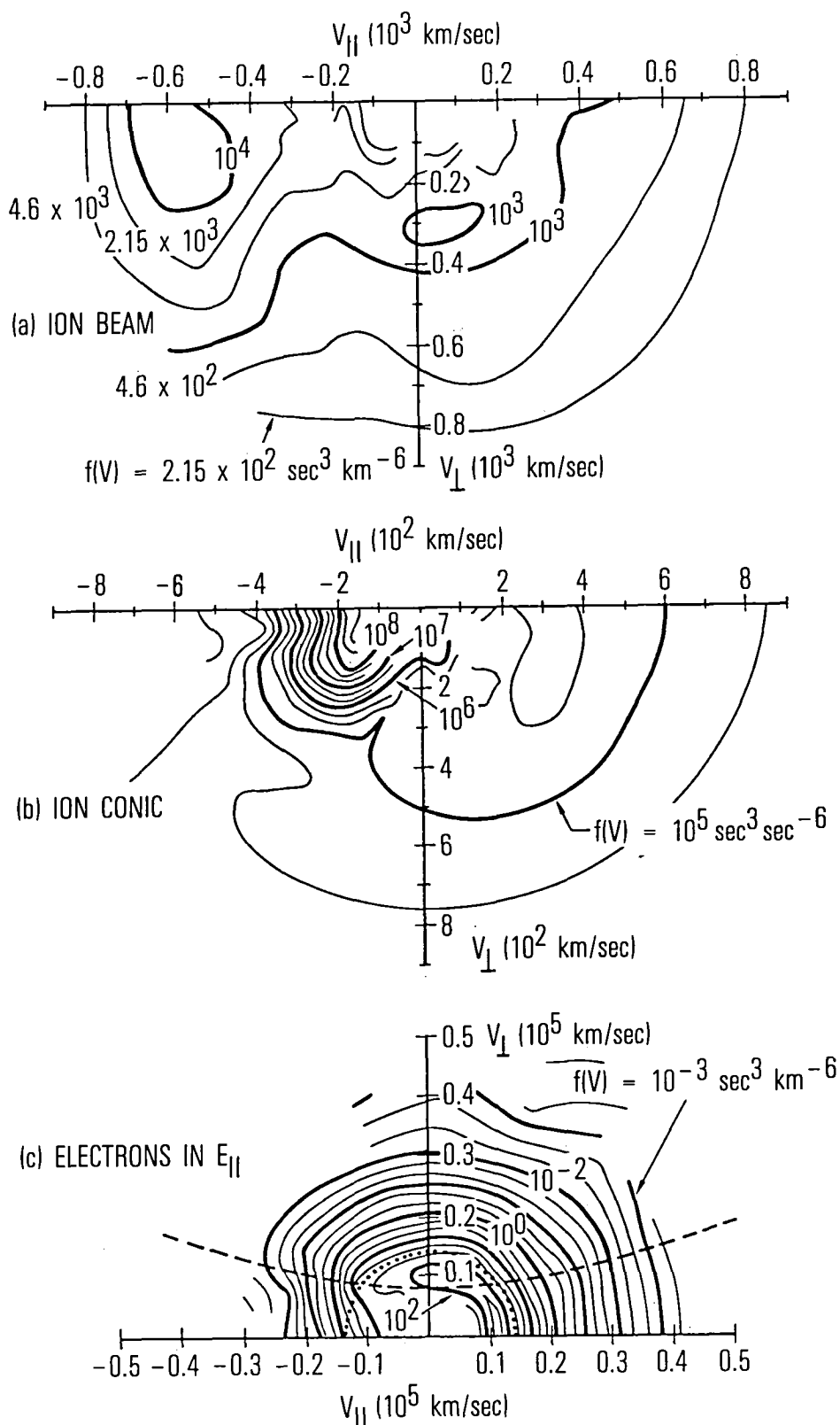


Figure 7-9: Typical distribution function contours of (a) ion beam; (b) ion conic; and (c) electrons in an inverted-V structure.

In addition to effects by ionospheric plasmas on magnetospheric processes, the upper atmosphere is the source for populating the geocorona. The very recent first survey of the geocorona has shown both an asymmetry, with an antisunward tail, and a decrease in density in the region of the ring current during a magnetic storm (Rairden et al., 1983). Since charge exchange between ring current ions and the thermal hydrogen appears to be the dominant loss mechanism for the ring current, atmosphere-magnetosphere coupling processes also must be investigated (Williams, 1983).

Given that ionospheric ions are drawn out of the ionosphere and accelerated, mechanisms for trapping these particles and distributing them throughout the magnetosphere are still being sought. They probably alter the local potential structures and may affect the global potential distribution which, in turn, would modify magnetospheric convection. The resulting distributions of both the cold plasma and hot plasma will certainly modify the location and strength of hydromagnetic wave resonances in the magnetosphere-ionosphere cavity (Greenwald, 1982).

All these observations indicate processes at work in the interaction region between the cold ionospheric plasma and the hot magnetospheric plasma. What the details of these additional processes are is far from clear at present (Hultqvist, 1982). It remains to be determined when, where and why different processes work and their relative contributions as source processes for different magnetospheric plasma regions.

### III. INTERMEDIATE SCALE COUPLING

In this section, we consider plasma and field phenomena which are specific to some sector of the magnetosphere and ionosphere.

#### EVENING POTENTIAL STRUCTURE

The evening discrete arc, distinguished by its optical intensity and well-defined morphology, has been a major focus of solar-terrestrial research. Early auroral electron observations (e.g., Frank and Ackerson, 1971; Gurnett and Frank, 1973; Evans, 1974) revealed distinctive signatures in their distribution. Satellite observations in the formation region of these auroral electron features have demonstrated a coherent picture in which auroral electrons are accelerated by an electric potential structure with an electric field component aligned with the magnetic field (e.g., Mizera et al., 1981; Fennell et al., 1981). Figure 7-10 shows a schematic illustration of the simultaneous occurrence of these plasma and electric field signatures and their spatial characteristics. The coherence of the data makes it rather difficult to argue against the qualitative interpretation that they signify the existence of parallel electric fields. Quantitative evidence for the existence of the auroral electric potential structure is available in the detailed adiabatic features and demarcations of plasma distribution functions (Mizera and Fennell, 1977; Chiu and Schulz, 1978; see review in Chiu et al., 1983a). Figure 7-9c shows the prominence of adiabatic demarcations in auroral electron distribution features. The adiabatic signatures in these plasma distributions is of great importance to arc formation theory; perhaps as important as the notion of the parallel electric field itself.

Another important point to be made is that  $E_{\parallel}$  alone is not sufficient to provide us with an arc formation mechanism, for an arc is an arc because it has a width. Figure 7-10 shows that there is a spectrum of scales transverse to the magnetic field in arc phenomena. By the same token, a spectrum of scales parallel to the magnetic field is almost certainly to be expected. The simplest argument against a single potential drop of  $\leq 100$  Debye lengths in parallel scale is that such a

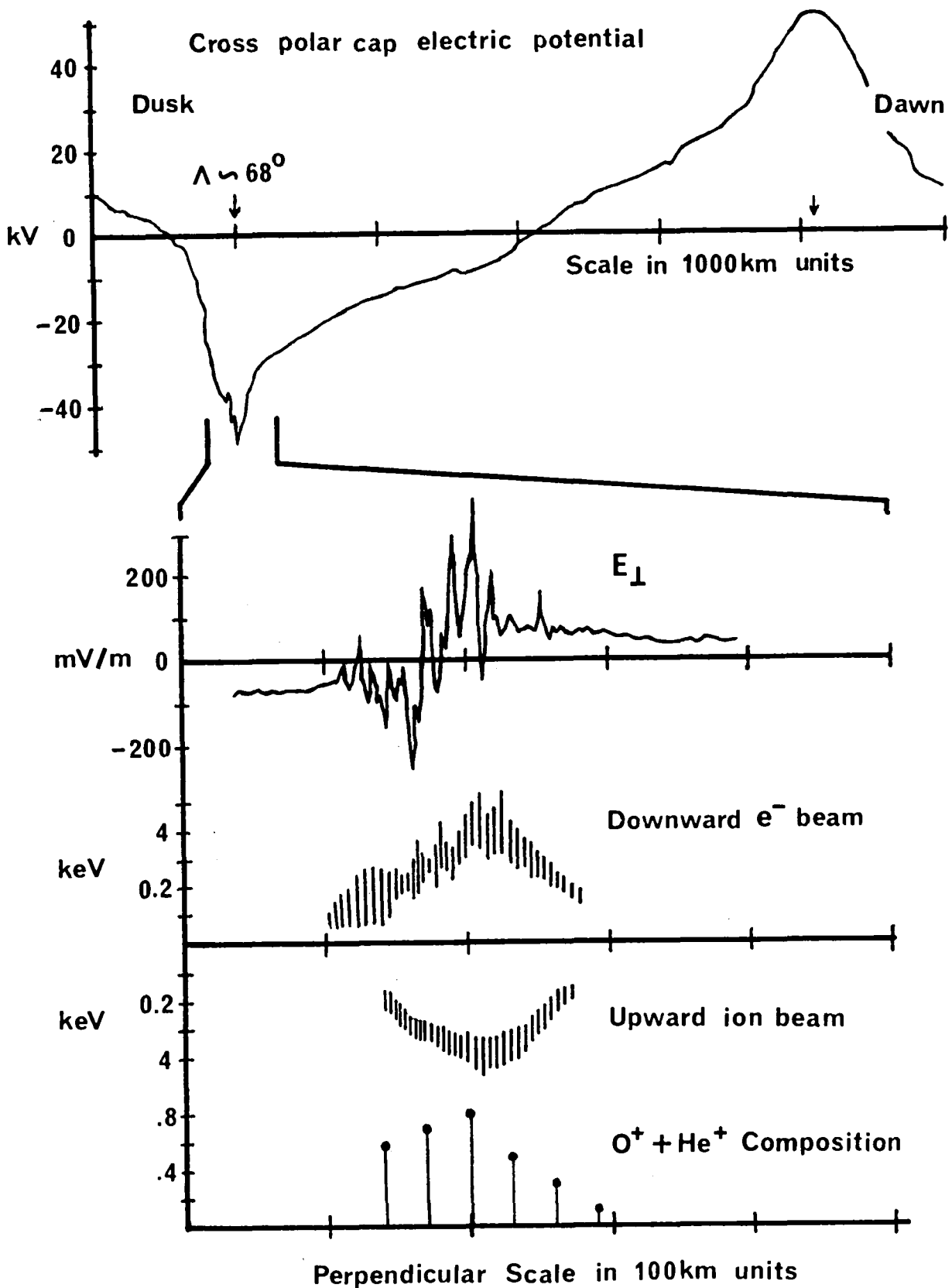


Figure 7-10: Observational features of electric field and plasma associated with a discrete arc.

structure cannot supply ionospheric  $O^+$  (e.g., Shelley et al., 1976) to the ion beam unless it is located near the ionosphere, in which case, it cannot simultaneously supply a downward electron beam at  $\sim 1 R_E$  altitude. However, this is not to say that the total potential drop cannot be made up of numerous discrete steps of small magnitude (Temerin et al., 1982). Thus, the field-aligned structure of the electric potential, as revealed by data to-date, is somewhere between the analogy of a ramp and a staircase of several thousand kilometers total length. For this reason, we have chosen the nomenclature "auroral potential structure" to denote the totality of arc formation structures, instead of "double layers" or "electrostatic shocks" which carry with them, perhaps unjustifiably, theoretical notions of scales assumed prior to observations.

As observations developed, theoretical alternatives on arc formation have been narrowed down. We are now entering the stage of quantitative comparisons between theory and experiment, a stage rarely seen in solar-terrestrial physics. An important quantitative comparison has been based upon the current-voltage relation along auroral field lines and upon the analysis of plasma distribution functions measured by auroral satellites and rockets.

Two particle populations can contribute to a field-aligned current: the ionospheric plasma and the high-altitude plasma in the plasma sheet or magnetosheath. The ionospheric plasma can supply only an upward field-aligned current of magnitude  $\leq 10^{-7}$  to  $10^{-6}$  A/m<sup>2</sup> (Lennartsson, 1980; Lyons, 1980), while field-aligned currents associated with auroral are upwards and generally from  $10^{-6}$  A/m<sup>2</sup> to a few times  $10^{-5}$  A/m<sup>2</sup> (Kamide and Rostoker, 1977; Anderson, 1978, and references therein). Recent observations (Burke et al., 1983; Sugiura et al., 1983) have indicated upward currents of  $>10^{-4}$  A/m<sup>2</sup>. Thus the upward current supplied by the ionospheric plasma can generally be neglected over the aurora.

If one assumes the plasma in the auroral flux tube to be non-diffusive, a kinetic relationship between field-aligned current  $j_{\parallel}$  and potential drop along the tube  $V_{\parallel}$  can be derived (Knight, 1973; Lemaire and Scherer, 1974; Chiu and Schulz, 1978). Such a relationship has been verified observationally (Lyons et al., 1979; Fridman and Lemaire, 1980; Lyons, 1981b).

The generation of the large-scale "inverted-V" precipitation regions by spatial variations in the magnetospheric convection electric field with  $\nabla \cdot E < 0$  has been investigated by Lennartsson (1977); Lyons (1980); and Chiu and Cornwall (1980). A commonly observed example of such an electric field divergence exists along the evening boundary between antisunward and sunward convection (Figure 7-1), a boundary along which "inverted V" precipitation regions are observed (Frank and Gurnett, 1971; Gurnett and Frank, 1973). Other variations in the electric field with  $\nabla \cdot E < 0$  also exist in association with the discrete aurora and over the polar caps in association with polar cap auroras (Chiu and Gorney, 1983).

By solving the equation for current continuity in the ionosphere, the currents, field-aligned potential differences, and particle acceleration associated with large-scale ( $\sim 100$  km in latitudinal width) "inverted-V" precipitation regions were quantitatively related. A negative divergence of the high-altitude, magnetospheric electric field is required. However, the "inverted-V" scale size was found to be a natural consequence of the current-voltage relations along auroral field lines and in the ionosphere, so the scale size need not be imposed by structure in the high-altitude electric field.

Smaller-scale auroral structure such as that associated with discrete auroral arcs ( $\lesssim 10$ s of km in latitudinal width) cannot be explained as a natural consequence of the current-voltage relationships without imposing structure in the high-altitude electric field distribution (Lyons, 1981a; Chiu et al., 1981a). The required structure has been inferred by Lyons (1981a), using observations of electric fields and precipitating electrons from the POPAR 3 rocket flight (Maynard et al., 1977; Evans et al., 1977). The measured and calculated precipitating electron energy fluxes are compared in Figure 7-11 and the agreement can be seen to be quite good.

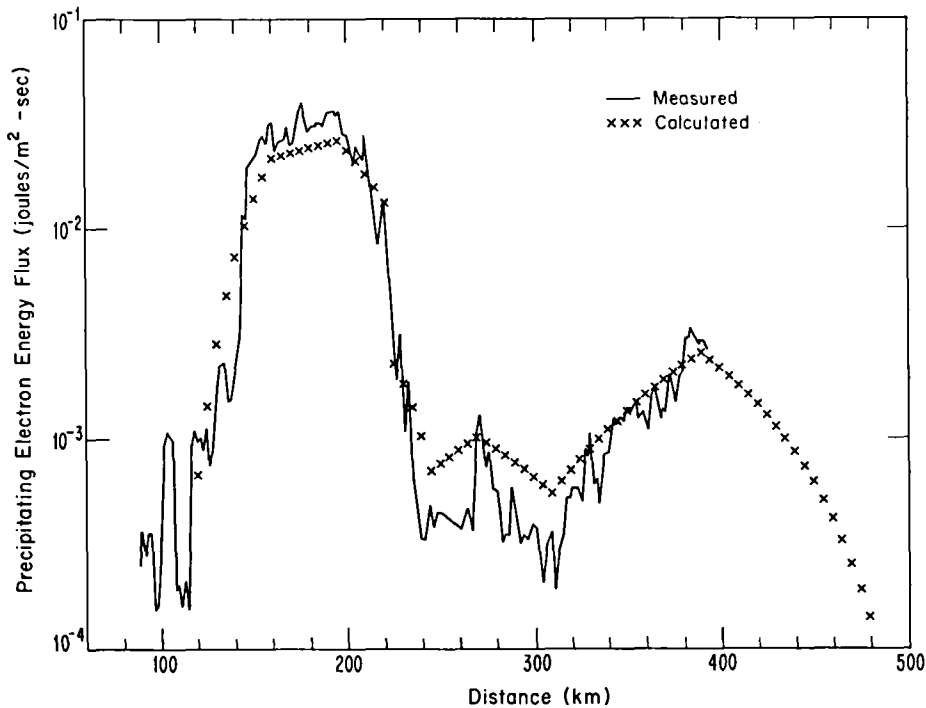


Figure 7-11: Measured precipitating electron energy flux from the POLAR 3 flight and the calculated precipitating electron energy flux from the solution to the ionosphere current continuity equation (from Lyons, 1981a).

A second area of quantitative comparison between theory and experiment attempts to move away from self-consistency between particle and electric field measurements. This approach argues that excess charges precipitated in the aurora must be returned to the magnetosphere via a return current; thus, completion of the circuit would result in a testable relationship between the parallel potential drop and the structure of the perpendicular electric field (Chiu et al., 1981a). Such a relationship is in agreement with constructions of the potential structure from electric field data (Temerin et al., 1981; Torbert and Mozer, 1978) on inverted-V scales.

A complementary approach to arc formation theory focuses on plasma turbulence (Hudson and Mozer, 1978) which is invariably associated with the adiabatic signatures discussed above. Recent work by Lysak and Dum (1983) suggests that arc-scale ( $\ll 10$  km) structures may be amenable to this approach.

Auroral plasma exhibits characteristics which indicate that it has gone through adiabatic auroral processes. However, at the same time, the existence of turbulent electric field and diffusive signatures (indicating some measure of turbulence) in auroral plasmas cannot be denied (see review in Chiu et al., 1983a). The coexistence of both adiabatic and diffuse signatures must be recognized in all theories of auroral phenomena (Chiu et al., 1981b). Insofar as the magnetosphere-ionosphere coupling processes (which give rise to the plasma signatures themselves) are concerned, there are two complementary points of view: (1) the turbulence is driven by non-Maxwellian distributions formed by adiabatic auroral processes, and (2) the adiabatic auroral structures are the consequences of turbulent processes. These alternatives are the basis of alternative view points in arc formation theory.

Chiu and Cornwall (1980) constructed a theory of a quiet arc in which plasma turbulence is assumed not to play a role. It is argued that adiabatic distributions (Chiu and Schulz, 1978) can support an electrostatic potential structure corresponding to the inverted-V scale. However, it was pointed out subsequently (Chiu et al., 1981b; Cornwall and Chiu, 1982) that a measure of turbulence must accompany the theory because the adiabatic distribution functions are expected to be unstable to a number of plasma instabilities which would have the tendency to destroy the inverted-V scale structure.

Lysak and Dum (1983) recently examined the dynamic response of an auroral flux tube to a (two-fluid) MHD perturbation. The innovation of their model is the inclusion of currents related to fixed transport coefficients which couple the macroscale MHD equations to microscopic turbulence generated by the MHD currents. They found that (kinetic) Alfvén waves, which carry the magnetic field line response to an equatorial perturbation (current or voltage driven), reflect off of a region of turbulence which is generated by the field-aligned current carried by the Alfvén wave. In the reflection process, an Alfvén wave, bouncing between the equator and the turbulence layer at about  $1 R_E$ , forms an electrostatic structure in which the parallel electric field is supported by (EIC) anomalous resistivity and in which the perpendicular electric field is intensified. The structures resemble the electrostatic shocks described in the next section. Lysak and Carlson (1981) have shown that the Alfvén wave-turbulence coupling mechanism gives scale sizes consistent with discrete arcs while the kinetic equilibrium models (Chiu and Cornwall, 1980) give larger scale sizes consistent with inverted-V's.

The above two approaches, starting from opposite poles of the spectrum of theoretical opinion on arc formation, are actually complementary because both arguments, when developed, must accommodate a spectrum of auroral scales with adiabatic processes prominent in the largest scales and turbulent processes responsible for small arc-structure scales. Thus, theoretical investigations of the last few years have probably dispelled the notion of magnetosphere-ionosphere coupling as a single monolithic process; rather, the current picture seems to be a coupling taking place over a spectrum of scales which are interrelated through energy cascade to energy sources at the plasma sheet.

It is by now clear that the discrete aurora is a complex phenomenon, encompassing a spectrum of temporal and spatial scales of the accelerating potential. As such, there is as yet no single mechanism which accounts for all scales, thus, the wave turbulence approach, with observational support from small-amplitude electric-field structures, is actually complementary to the adiabatic approach which finds observational support in the inverted-V scales. Since these two approaches probably bound the auroral scale spectrum, studies on the interaction between large and small scale phenomena (e.g., Cornwall and Chiu, 1982) may perhaps point the way towards a unified theory of the auroral potential structure.

## ELECTRIC FIELD SIGNATURES

As has been demonstrated in the subsection on auroral acceleration potential structure above, plasma and field structures cannot be separately discussed in MIC processes. However, observations of electric fields in the last few years have revealed such a wealth of new features, not just confined to the discrete-arc potential structure, that it is worthwhile to pay special attention to them. The spectrum of MIC scales is most clearly illustrated in the electric field data.

Here we refer to the observational distinction between DC and AC electric field signatures based on time scale in the data as opposed to theoretical arguments about the origin of both. The electrostatic shocks (somewhat loosely defined here as a sharp change of electric field amplitude) observed by Mozer et al. (1977) on the 0.5-5 second time scale (see Figure 7-12) will be

referred to as DC, while the data in Figure 7-13 on the millisecond timescale on the same satellite in the same spatial region ( $1 R_E$  auroral field lines) will be referred to as AC. The electrostatic shock data have been correlated with discrete arcs (Torbert and Mozer, 1978; Kleitzing et al., 1983) which sometimes persist for hours in the auroral ionosphere, and with VLF saucer events in the wave data (Temerin, 1981) which also persists for entire auroral zone crossings of the satellite. This relationship is related to that schematically illustrated for the auroral potential structure in Figure 7-10.

The electrostatic shocks in Figure 7-12 occur primarily in the perpendicular electric field component with amplitudes typically in the range 0.1-0.5 V/m. They often occur in pairs, which is consistent with a V-shaped equipotential profile. In such a case, one must distinguish between the width of the intense DC electric field region at the edge and the width of the "V-shock" as a whole. The former is typically a few kilometers, while the latter can be the order of the inverted-V scale size (100 km). There is, however, very little correlation between the two scales. Often, only a single or "S-shock" event is observed. Construction of the potential structure shows that a parallel electric field must be associated with these shocks (Torbert and Mozer, 1978).

The AC electric field data in Figure 7-13 show some very distinct phenomena on the millisecond timescale. The data are typical of an upward flowing ion beam region above much of the auroral potential drop. Electrostatic ion cyclotron (EIC) waves are apparent in the perpendicular electric field component, along with a low frequency oscillation ( $\sim 10$  Hz) identified as low frequency turbulence (Temerin, 1978). The latter might be considered DC if interpreted as a temporally constant spatially varying electric field, through which the spacecraft flies. Since its source has not yet been identified conclusively, whether it is classified as AC or DC depends on where one draws the timescale line. It is a pervasive phenomenon throughout the auroral zone, cusp, and into the polar cap at all altitudes where real-time wave data are available.

A great deal of work has been done on the EIC instability in terms of its effect on particle heating and plasma transport properties. It is still unclear whether the electron current or the ion beam, which always accompanies it, is the source of free energy for the linear instability. The large amplitude, ( $e\phi/T_e \sim 1$ ) and coherence of the observed waveform in Figure 7-13 have had a major impact on nonlinear theories of the EIC instability and its effect on particle heating (see "Ion Conics and Beams" in Section IV).

Focusing now on the parallel component in Figure 7-13, one sees symmetric and asymmetric structures on the millisecond timescale which are not well correlated with the periodic waves in the perpendicular component, except that they occur in the same region of space. These have been identified by Temerin et al. (1982) as solitary waves and double layers, respectively. The latter, thought to evolve from the former in the presence of an electron current or ion beam, are potentially significant to the auroral acceleration process because of their average net upward electric field, particularly evident in panel (b). Each double layer has an amplitude  $e\phi/T_e \lesssim 1$ , but in aggregate they can contribute substantially to auroral acceleration. They are observed about 5% of the time in Figure 7-13 with amplitudes around 10 mV/m, yielding in average electric field of 0.5 mV/m, or 1 kV over 2000 km.

An alternative nonlinear model of anomalous resistivity has been developed as a possible explanation of parallel electric fields on auroral field lines (Rowland et al., 1981; Rowland and Palmadesso, 1983). This model predicts parallel potential drops in the presence of both ion density fluctuations and a field-aligned current which exceeds a threshold value. The ion density fluctuations must have large amplitude ( $\delta n_i/n_i \gtrsim 2$ ) but may be generated by any one of several possible processes. The characteristic profile or signature of the resulting potential structure

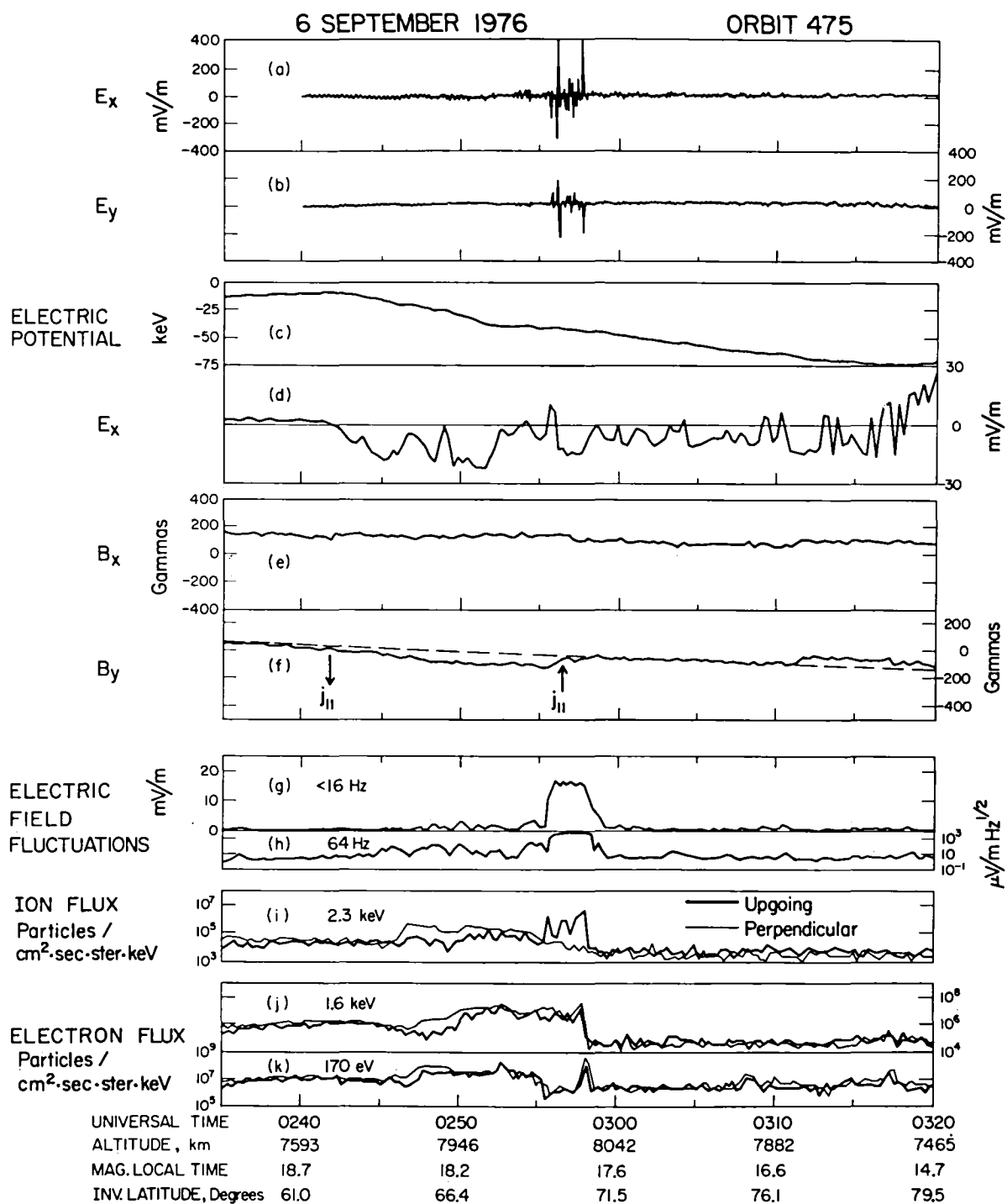


Figure 7-12: Electrostatic shocks, the large scale electric field, field-aligned currents, waves and particles measured during a typical S3-3 satellite passage from the plasmasphere to the polar cap (Mozier et al., 1977).



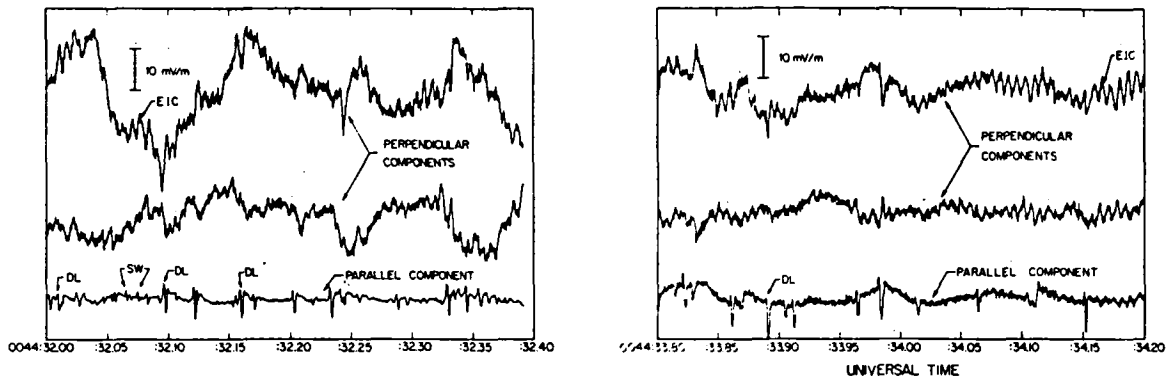


Figure 7-13: The two perpendicular and one parallel electric field components shown. Examples of double layers (DL), solitary waves (SW) and electrostatic ion cyclotron (EIC) waves are marked. These data were acquired on August 11, 1976, at an altitude of 6030 km, an invariant latitude of  $74.1^\circ$ , and a magnetic local time of 15.74h. The finite parallel component of the low-frequency turbulence (i.e.,  $\sim 10$  Hz) on the right side of the figure is an artifact due to frequency-dependent gain of the detectors at low frequencies (Temerin et al., 1982).

depends on both the ambient transverse or convection electric field which is not directly affected by the resistivity and on the nature of the ion wave turbulence which initiates the resistivity generation process.

In general, a field-aligned current with a finite transverse extent and smooth gradients in the direction transverse to the magnetic field would be expected to produce a "V" shaped average potential structure in the absence of a strong transverse electric field and an "S" shaped structure if  $E_\perp$  is large. In addition, one would expect a fluctuating electric field component associated with the plasma waves which sustain the resistivity. The signature of the parallel electric field characteristic of this anomalous resistivity model depends critically on the nature of the ion density fluctuations present. Computer simulations of this mechanism conducted by Rowland and Palmadesso (1983) show the presence of potential drops which are somewhat steeper and more localized than the characteristic scale length for variations of the ion density along the magnetic field ( $\sim 2\pi/k_\parallel$  for an ion wave). In the case of EIC waves which propagate nearly perpendicular to the magnetic field, one would expect parallel potential drops with a characteristic spacing  $\sim 10$  times the local EIC wavelength ( $2\pi/k$ ) and a scale size which may be a small multiple of  $2\pi/k$ . Ion density "cavitons" produced by the oscillating two stream instability (Freund et al., 1980) are also capable of supporting anomalous resistivity. The cavitons are very highly localized ( $\sim$  tens of Debye lengths in size and separated by distances which may be 10-100 times this value) and lead to highly localized potential drops with similar size and spacing. This last case is in contrast to the more uniform electric field structure expected on the basis of earlier theories of anomalous resistivity and is consistent with the often spiky nature of the parallel electric field component observed in satellite data.

In addition to the shape of the potential structure, one can make favorable comparisons between the characteristic shape of the electron velocity distribution function seen in particle simulations (see Figure 7-14) of the resistivity mechanism (Rowland et al., 1981) and electron velocity distributions observed in "inverted-V" events (e.g., Arnoldy, 1981).

Perhaps surprisingly, independent attempts to find support in auroral observational data for theoretical models of field-aligned potential drops based on double layers, electrostatic shocks, kinetic Alfvén waves, anomalous resistivity, and adiabatic processes have all had some success. As a result, there is an emerging consensus that the electrodynamics of ionosphere-magnetosphere coupling involves a complex interplay between several, if not all, of these mechanisms (Chiu et al., 1981b). Additional studies aimed at clarifying the way in which different electrodynamic processes interact and at identifying the regions of parameter space in which individual processes might dominate are much needed.

### MORNINGSIDE STRUCTURES

When viewed with low optical resolution, the auroral luminosity takes the characteristic oval shape. However, upon closer examination, drastic differences between the eveningside and the morningside of the auroral oval begin to appear. Indeed, differences in the plasma signatures in the arc formation region (>5000 km altitude) would argue for completely different magnetosphere-ionosphere coupling processes for the two sides of the oval, a major unsolved question in the global view of the coupling.

Ground-based observations have long characterized auroral morphology into evening discrete arcs and morningside multiple arc systems and patches with tendency for pulsation (Oguti, 1981; Royrvik and Davis, 1977; Johnstone, 1978). A schematic summary (Figure 7-1c) of such auroral morphological features has been given by Akasofu and Kan (1980). However, since optical configurations are only an imperfect indicator for arc formation dynamics, it is necessary to consider these local time dependences in terms of their plasma and electric field signatures at arc formation altitudes, as in the case of evening discrete arcs. Figure 7-15 shows the distinct differences of plasma signatures between the evening and the morning sectors (Chiu et al., 1983b). In this symmetrical pass of the S3-3 satellite across the auroral oval, it is seen that the simultaneous occurrence of an electron inverted-V and an ion beam indicates parallel potential drops in the evening sector both above and below the satellite; while, on the morningside, the absence of both features in the plasma observations is evidence that the morningside aurora is not associated with parallel electric potential structures, either above or below the satellite. These features of Figure 7-15 are clearly supported by statistical preference of downward electron beams and upward ion beams for the evening sector (Hoffman and Lin, 1981; Gorney et al., 1981).

If the parallel electric field plays no role in morningside arc formation, then two major questions in the global structure of magnetosphere-ionosphere coupling must be addressed:

- What is the arc formation mechanism in the morning sector?
- How does this new mechanism fit into the global coupling and into the global structures of magnetospheric currents and convection?

At present, neither of the questions have been investigated to any satisfactory degree. An attempt to address the first question in terms of wave particle interactions has been made (Chiu et al., 1983b). These authors invoke intrinsic differences in the pitch angle distributions of ions convected to the two sectors as the source of local time asymmetry, favoring the generation of mirror instability on the morningside and giving rise to modulation of auroral electron precipitation there.

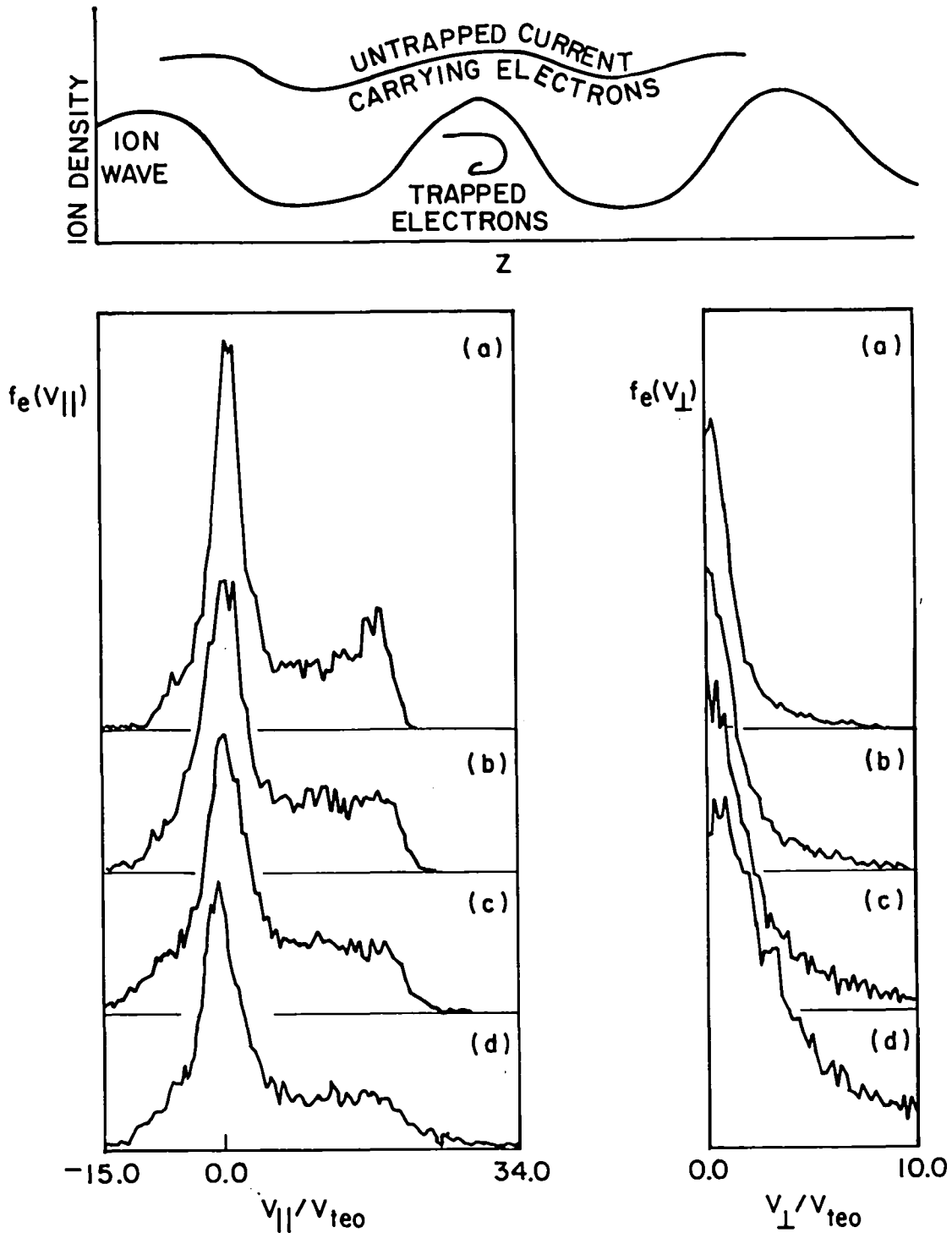


Figure 7-14: In the resistivity model of Rowland et al. (1981), current carrying electrons interact with lower-energy electrons trapped in potential wells associated with large-amplitude ion density fluctuations. The resulting momentum transfer provides resistivity. Characteristic electron distribution functions in a computer simulation are shown for various times: (a)  $2400/\omega_{pe}$  (b)  $3200/\omega_{pe}$  (c)  $4800/\omega_{pe}$  (d)  $8800/\omega_{pe}$ .

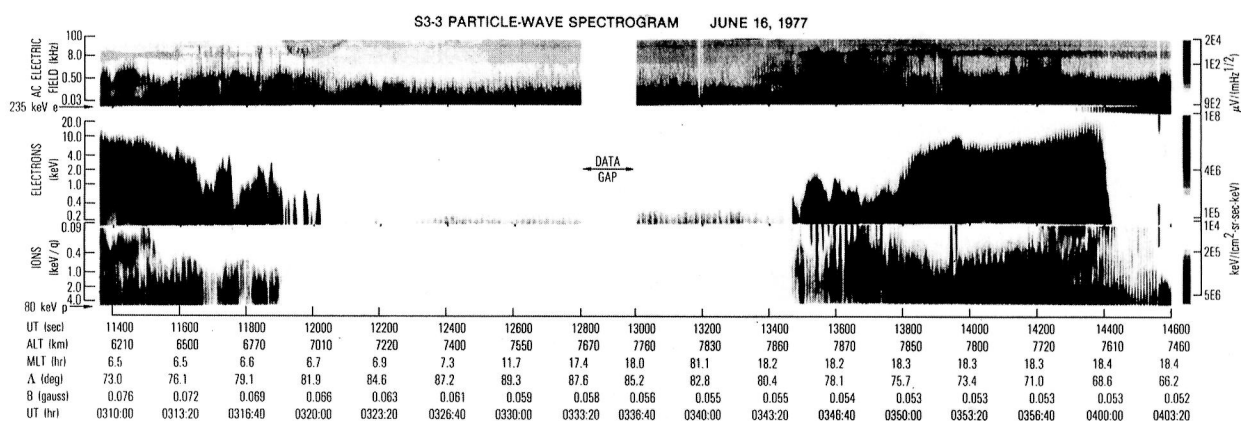


Figure 7-15: Energy-time spectra of a dawn-dusk pass of the S3-3 satellite. The top panel is a wave frequency-time plot of the pass.

## WESTWARD TRAVELING SURGE

An example of intermediate-scale MIC processes is the westward traveling surge. The intense field-aligned current associated with the surge itself is carried by field-aligned electron beams which have been electrostatically accelerated at great altitudes. The associated current sheets stretch out to beyond 10,000 km altitude and close through ionospheric currents. The intense current sheet drives the magnetospheric plasma unstable to waves that heat the ions to a couple of hundred eV in a fraction of a minute.

The ions are heated strictly transverse to the magnetic field in these events and as a result they experience a strongly increased mirror force which moves them up the magnetic field lines. In an event observed by the ISEE-1 ion mass spectrometer at 10,500 km altitude, the transversely heated plasma from the altitude range 3,000 to 10,500 km, mainly consisting of oxygen, was moved out past the satellite in a period of about fifteen minutes (E. Ungstrup, private communications, 1983).

All the anisotropically heated plasma, left behind by the surge, will have to redistribute itself after its passage and it can be estimated that it will take a 200 eV  $O^+$  ion about one hour to reach the equatorial plane. Thus the redistribution of plasma takes about one hour and it is interesting to compare this redistribution time to the duration of substorms as observed by ground magnetometers. The volume of the magnetosphere affected by a westward traveling surge is the volume connected to an area of the Earth that is  $60^\circ$  in longitude and  $5^\circ$  to  $10^\circ$  in latitude along the auroral oval. The plasma motions in this large volume give rise to currents in both the magnetosphere and ionosphere during the substorm surge event.

The westward traveling surge is perhaps one of the most spectacularly intense events of MIC, but its dynamics has yet to be understood. The observational scenario described here represents an attempt to organize separate pieces of data into a coherent picture of this interesting phenomenon.

#### IV. MICROSCOPIC SCALE EFFECTS

A large number of different types of plasma waves, including whistlers, electrostatic ion cyclotron waves, ion acoustic waves, lower hybrid waves, Langmuir waves, etc., have been observed and/or theoretically predicted to occur on auroral field lines (e.g., Gurnett and Frank, 1977; Kindel and Kennel, 1971; Papadopoulos and Palmadesso, 1976; Temerin, 1981), and there exists a correspondingly large variety of wave-particle-interaction phenomena in the ionosphere-magnetosphere coupling regions. These processes, as a group, are important for several reasons. They can, for example, give rise directly to phenomena of intrinsic interest, such as auroral kilometric radiation which appears to involve a relatively efficient conversion of particle energy to electromagnetic energy. They can give rise to phenomena valuable because they provide means of diagnosing space plasmas, as in the case of plasma density estimates based on remote observation from the ground of ionospheric ion acoustic waves or Langmuir waves ("plasma lines"). Finally, wave-particle interactions can play an important role in the dynamics of I-M coupling phenomena at small scale sizes and, less directly and to a somewhat lesser extent, at larger scale sizes as well. The impact of wave-particle interaction phenomena on large scale dynamics arises as a result of modification of the plasma transport coefficients through the action of plasma turbulence.

#### AURORAL KILOMETRIC RADIATION

Auroral kilometric radiation (AKR) is the most intense terrestrial radio emission associated with the coupling of the magnetosphere and the ionosphere. Numerous studies of satellite observations over the past decade have now provided us with detailed characteristics of AKR and associated wave-particle interactions. Results of these studies have had a large impact on AKR generation theories and on theories regarding potential structure and particle distribution functions in the auroral region.

Recent detailed reviews of the observational results may be found in Gurnett (1974), Gurnett and Anderson (1981), Green (1981), and Benson (1981). AKR is an intense electromagnetic radio emission generated at frequencies from 30 to 700 kHz as bands of noise usually with sharp upper and lower cutoff frequencies and with peak intensity in the few hundred kHz frequency range (Figure 7-16). The intensity of AKR is highly variable, often changing by as much as 60 to 80 dB on time scales of ten minutes or less. The total AKR power radiated from the Earth can be as high at  $10^9$  W, although a more typical average value is about  $10^7$ - $10^8$  W. The radiation is emitted in a conical emission pattern, the axis of which is tipped away from the Sun toward about 21 to 24 hours magnetic local time (Gurnett, 1974; Calvert, 1981b). Most studies have concluded that AKR is generated in the right-hand R-X mode (Gurnett and Green, 1978; Kaiser et al., 1978; Benson and Calvert, 1979; Shawhan and Gurnett, 1982). It is possible that AKR may consist of a mixture of both R-X and left-handed L-O (Oya and Morioka, 1982) modes but most observations indicate that the dominant component is right-handed. A second harmonic component of AKR is occasionally observed and is usually weaker than the fundamental (Benson, 1982; Anderson, 1983; Shawhan, 1983). In the DE-1 data, Shawhan (1983) has found evidence of the fundamental being right-handed and the weaker second harmonic being left-handed. AKR has been shown to

be associated with discrete auroral arcs (Gurnett, 1974) and field-aligned currents (Green et al., 1982). Gurnett (1974) suggested that AKR is generated by intense "inverted-V" electron precipitation bands and this has been confirmed by direct *in situ* measurements in the auroral regions (Benson and Calvert, 1979; Green et al., 1979; Benson et al., 1980).

Much effort has gone into comparing the numerous theories of AKR generation with the actual wave and particle observations. Because of the high efficiency ( $\sim 1\%$ ) with which AKR is generated, Gurnett (1974) concluded that AKR is probably generated by a coherent plasma instability rather than by incoherent radiation from energetic electrons. Source region observations of the waves and particles associated with AKR are beginning to restrict the possible theories of AKR generation.

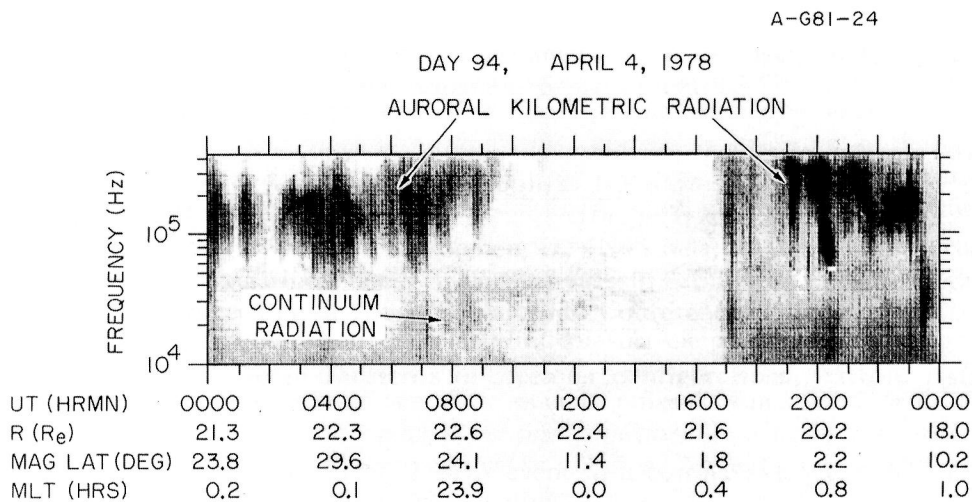


Figure 7-16: A 24-hour ISEE 1 sweep frequency receiver spectrogram from 10 kHz to 400 kHz of auroral kilometric radiation observed beyond  $18 R_E$  near midnight. Note the upper and lower cutoff frequencies and intervals when the radiation disappears completely (from Gurnett and Anderson, 1981).

A number of AKR theories have been proposed (Benson, 1975; Barbosa, 1976; Roux and Pellat, 1979; Melrose, 1976), but further observations are needed to see if the special requirements of these theories are met. Wu and Lee (1979) suggested that relativistic effects should be included in the Doppler-shifted cyclotron resonance and that a loss cone region in the electron distribution would be unstable and give rise to AKR. They assumed that a field-aligned potential drop would not only produce an enhanced upgoing loss cone but would also deplete low-energy ionospheric particles and form density cavities which satisfy their  $f_p < f_g$  condition for the generation of AKR. Mizera and Fennell (1977) have reported enhanced loss cones in the auroral regions presumed to be due to large parallel electric fields. It has also been shown that AKR originates within local depletions of electron density (Benson and Calvert, 1979; Benson, 1981; Calvert, 1981a). Growth rate calculations, based on measured particle distribution functions in the auroral regions and using the relativistic Doppler-shifted electron cyclotron resonance condition, have yielded growth lengths for 10 e-foldings of the order of the typical North-South thickness of the inverted-V region ( $< 200$  km) (Omidi and Gurnett, 1982; Dusenbery and Lyons, 1982; Melrose et al., 1982).

The predominant view held by most researchers is that AKR emissions are generated by Doppler-shifted cyclotron radiation from the electrons involved in the auroral acceleration process. As the generation mechanism(s) become better understood, we can learn more about the auroral acceleration processes from the detailed characteristics of the AKR emissions. Gurnett and Anderson (1981) (assuming AKR generation by Doppler-shifted cyclotron radiation) showed that the observed cutoffs were controlled: (1) by the plasma density distribution; and (2) the detailed profile of the electrostatic potential distribution. Under certain conditions, the low-frequency cutoff gives a direct indication of the upper altitude limit of the acceleration region.

An exciting recent discovery was that AKR contained much fine structure. Gurnett et al. (1979) and Gurnett and Anderson (1981), using wideband analog data from ISEE, showed that AKR has a very complex frequency-time structure; a typical AKR wideband spectrogram is shown in Figure 7-17. Narrow bandwidth bursts with rapidly drifting center frequencies were commonly observed. Some of the discrete features resembled whistler-mode triggered emissions, suggesting that the basic instability might be quite similar. If the AKR fine structure is a result of disturbances propagating through the generation region, the velocities inferred from the frequency drift rates are comparable to the ion acoustic speed in the generation region (Gurnett and Anderson, 1981). Grabbe et al. (1980) proposed that AKR could be generated by the beam amplification of electromagnetic waves via coherent electrostatic ion cyclotron density fluctuations. A few rare observations of ion cyclotron harmonic structure in AKR have been observed (Grabbe, 1982). Calvert (1982) has proposed a feedback model for the generation of AKR in which the frequency of the radiation is dependent on an integral number of wavelengths fitting into local density enhancements at the source, much smaller than and embedded inside the cavity.

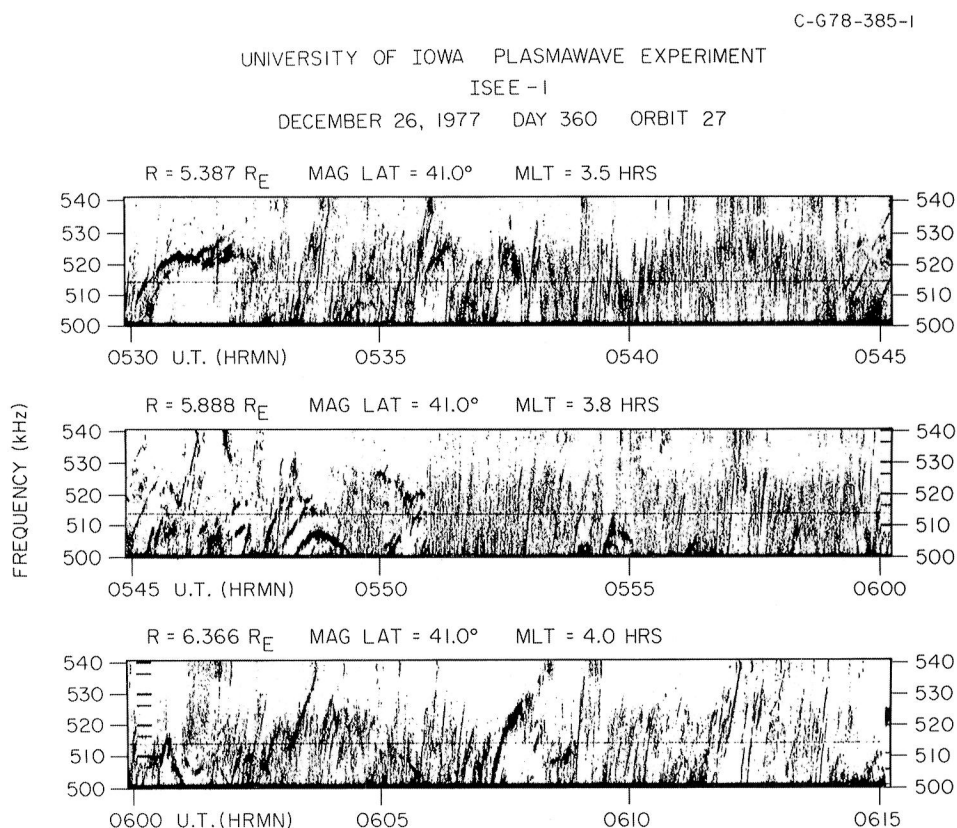


Figure 7-17: Dynamic spectra of auroral kilometric radiation in the 500-400 kHz range observed on ISEE 1 (from Gurnett et al., 1979).

## ION CONICS AND BEAMS

One of the major surprises in space plasma physics in the past decade is that ions are commonly accelerated perpendicular to the magnetic field to temperatures at least one hundred times larger than their initial temperatures (Figure 7-9b). Experimental evidence for perpendicular ion acceleration (ion conics) has been accumulated from at least four magnetospheric satellites and two sounding rockets (e.g., Shelley et al., 1976; Sharp et al., 1977; Klumpar, 1979b; Ungstrup et al., 1979; Gorney et al., 1981; Whalen et al., 1978). However, evidences for the proposed mechanisms that produce perpendicular ion acceleration remain unclear. The evidence that does exist suggest electrostatic (ES) ion cyclotron waves may be involved (Kintner et al., 1978) although this evidence is not conclusive.

Kintner and Gorney (1983) performed an exhaustive study of the S3-3 data set, in search of examples of perpendicular ion acceleration and simultaneous broadband plasma waves. During perpendicular ion acceleration events, plasma waves were identified at the lower hybrid resonance (LHR) frequency. The LHR wave amplitude was .5-1 mV/m over a 1 kHz bandwidth at ~1-2 kHz frequencies. Although the electrostatic  $O^+$  cyclotron mode was not explicitly identified in their data, a maximum broadband amplitude of 2-3 mV/m was inferred, assuming that all turbulence at frequencies below 500 Hz was due to Doppler-broadened  $O^+$  cyclotron waves. No evidence was found for the existence of electrostatic hydrogen cyclotron waves in the ion conic source region. Also, small scale DC potential jumps were not spatially well correlated with regions of perpendicular ion acceleration.

Gorney et al. (1982) observed an event in which energy from lower hybrid waves was being absorbed by energetic ions at harmonics of the ion cyclotron frequency. Had this energy absorption occurred throughout the low-altitude region at a constant rate, a total energy gain of 5 eV/ion was implied. This energy gain was insufficient to explain the several hundred eV ion conics which were observed.

While a great deal of attention on wave-particle interaction observations have been focused upon the ion conic, we must be reminded that conic heating should perhaps be discussed in the context of wave-particle heating of both ion beams and conics.

Although no conclusive evidence for the source of perpendicular ion acceleration exists, a variety of theories have been developed to explain ion conic acceleration. These theoretical mechanisms fall into three categories: acceleration by ES ion cyclotron waves (Palmadesso et al., 1974; Lysak et al., 1980; Papadopoulos et al., 1980; Dusenberry and Lyons, 1981; Singh et al., 1981; Ashour-Abdalla et al., 1981; Okuda and Ashour-Abdalla, 1981), acceleration by narrow oblique electric potential jumps (Yang and Kan, 1983) and acceleration by ES lower hybrid (LH) waves (Chang and Coppi, 1981). The free energy for either the ES plasma waves or potential jump must, of course, come from other sources such as field-aligned currents carried by either drifting thermal electrons or precipitating keV electrons.

Those theories which employ ES waves to transfer energy to the ions imply the production of a specific type of ion conic. Namely the accelerated ion distribution is formed nearly perpendicular to the magnetic field. After their initial acceleration the ions rise under the influence of the magnetic mirror force and form the more commonly observed ion conic with a folded (i.e., conical) distribution function. The hypothesis of nearly perpendicular acceleration is partially supported by the observation that ion conics are not seen traveling downward and that ion conics are seen with peak flux at  $90^\circ$  pitch angles (see Figure 7-9b). Ions accelerated by narrow oblique potential jumps should initially have pitch angles significantly away from  $90^\circ$ . This feature is not supported by the data.



Observations of particles and fields in the ion conic source region allow us to examine the consistency of several theories of ion conic generation. For example, Dusenberry and Lyons (1981), predict that ion cyclotron wave amplitudes of 10-20 mV/m can accelerate ions (quasi-linearly) up to 40 times their initial energy. According to this theory, the observed wave amplitude of 2-3 mV/m in the conic region is too small to account for the observed acceleration of 400 times (or more) of the initial ion energy. Lysak et al. (1980), by not making the random phase approximation requires coherent ion cyclotron waves with amplitudes of about 20 mV/m predict ion conic energies of 400 eV. Although large amplitude ion cyclotron waves are observed on auroral field lines, they are not observed in association with ion conics. For the LHR wave acceleration mechanism, Chang and Coppi (1981) assume an amplitude of 50 mV/m compared to the observed value as low as .5-1 mV/m and typical value of 5-10 mV/m. Similar points can be raised about present simulation studies of conic formation. Therefore, in general, the observed electric field amplitudes appear to be too small to satisfy the demands of present theories.

Both current and ion beam driven electrostatic instabilities will contribute to the parallel heating of the ion beam. The observation of ion conics in the electron beam region suggests that lower hybrid heating may be important there. Perpendicular EIC heating may be important at higher altitudes, in the ion beam region where those waves have been observed. However, perpendicular heating is masked by parallel acceleration in this region (Lysak et al., 1980). Thus, despite the extensive amount of work that has been done on perpendicular ion heating by EIC waves, there is still no solid experimental confirmation of this as the primary production mechanism for ion conics. Furthermore, a great deal has been said about preferential acceleration of ionospheric oxygen, partly because of the discovery that  $O^+$  constitutes a significant fraction of the magnetospheric plasma. However, only recently (ISEE, DE) have experiments had the mass, energy and pitch angle resolution to be definitive about the relative importance of  $O^+$  vs.  $H^+$  conics (Sharp et al., 1983), and it is not yet clear that preferential perpendicular acceleration of the larger  $M/Q$  species has been demonstrated.

The conics themselves may be both absorbers or emitters of wave energy around the lower hybrid frequency. Lower hybrid heating of the conics may be quasi-linear (Chang and Coppi, 1981), stochastic (Karney, 1978), or due to trapping (Lysak et al., 1980) while absorption of VLF wave energy is by cyclotron damping (Gorney et al., 1982). Conic emission near the lower hybrid frequency which can also produce gyroharmonic structure in wave spectrograms (Cattell and Hudson, 1982) requires  $\delta f / \delta v_{\perp} > 0$ , which is not a common feature in the particle data. Its absence may be due in part to quasi-linear diffusion associated with the wave emissions (Roth and Hudson, 1981). Until we have a better understanding of the ion conic production mechanism, it will be difficult to discern whether we are seeing a distribution function which has been nonlinearly stabilized, or was never unstable to begin with. Similar questions arise for the ion beam, where it is clear that the observed beams have been subjected to wave-particle heating. Therefore, other criteria must be found for distinguishing between possible free energy sources for observed waves, e.g., electron current as opposed to ion beams for generation of EIC waves, besides the stability of measured distribution functions. Sample criteria for comparison are frequency, wavelength, and bandwidth of observed waves against those predicted for a given source of free energy. As always, linear theory only sets a baseline for what wave modes are expected, as is evident from the variety of nonlinear waveforms in Figure 7-13.

Anomalous resistivity can lead to the selective acceleration of one or more ion species in a multi-ion plasma (H. Mitchell and P. Palmadesso, private communications, 1983). Low frequency electric fields maintained in a plasma by resistivity (either collisional or wave-induced) imply a balance between momentum transferred to the electrons by the electric field and that transferred by the resistivity from the electrons to other particle species. As a result, the electron fluid

does not experience a net acceleration due to the electric field though some "runaway" electrons may be accelerated. If the plasma contains only one ion species, the momentum balance is the same for the ions and the ion species is not accelerated either. If a second ion species is present which does not participate in the resistive momentum transfer, it may be accelerated by the electric field while the participating ions experience an acceleration opposite to the field. This mechanism may be applied to the auroral field line plasma in the presence of a current along the magnetic field. Significant resistivity between the electrons and hydrogen ions may accelerate the oxygen ions provided that the resistivity has the appropriate characteristics, i.e., provided that the resistivity is due to the interaction of the electrons with a plasma wave carried by the hydrogen ions such as a hydrogen EIC wave. This process would lead to beam-like oxygen flow as opposed to the "conic" oxygen flows thought to be produced by transverse heating of oxygen by oxygen EIC waves. Both types of oxygen flows have been observed (Horwitz, 1982).

Clearly we have a fundamental lack of understanding of the conic generation mechanisms. Whether the lack of understanding is theoretical, observational or conceptual remains to be seen. Perhaps present theories do not predict measurable quantities (such as saturation wave amplitudes) self-consistently, or perhaps observations have not yet been made exactly in the right place at the right time to adequately represent the ion conic source region. Ultimately, perhaps our concept of ion conic acceleration as a single-step microphysical process deserves scrutiny, since there seems to be an inability of present theories to yield satisfactory results.

It is clear that nonlinear wave-particle interactions and the plasma turbulence often associated with them play a critical role in ionosphere-magnetosphere coupling phenomena over a significant range of scale sizes. The complex interactions between these processes are only now beginning to be studied but it is clear that a fully satisfactory understanding of M-I coupling can be achieved only after all the major processes of this type have been identified and the major interactions between different microscopic processes and between microscopic and macroscopic processes have been elucidated.

## V. SUMMARY/CONCLUSIONS

Based on our assessment of progress in magnetosphere-ionosphere coupling (MIC) physics in the last five-six years, we envisage a period in the near future in which basic and global issues in MIC will be addressed observationally and theoretically. These expectations are derived from the following general conclusions of our assessment:

- Magnetosphere-ionosphere coupling (MIC) encompasses a complicated set of interacting phenomena occurring on vastly different time and spatial scales rather than a set of isolated processes. Magnetosphere-ionosphere coupling research should interrelate present and future observations of the diverse phenomena through theories and models to provide a coherent understanding of the magnetosphere-ionosphere system.
- Energy flow in the ionosphere-magnetosphere system is primarily from the magnetosphere to the ionosphere. However, it is now clear that the ionosphere is not a passive element in this system. The spatial distribution of ionospheric conductivity affects magnetospheric plasma transport and the configuration of the magnetospheric current systems. Atmospheric motions may significantly modify magnetospheric plasma circulation. Furthermore, a substantial portion of the plasma population of the magnetosphere originates in the ionosphere.

- Theory has developed considerably along with observation for electrodynamic coupling and related turbulent phenomena. However, the present state of theoretical development does not adequately address the feedback onto the magnetosphere from the dynamic ionospheric response to these processes. Currently, such theoretical analyses are just beginning but are required for complete understanding of magnetosphere-ionosphere coupling.
- The global scale theories of magnetosphere-ionosphere coupling have matured to the point that it is possible to test their assumptions and predictions. Thus, experiments can be designed to test specific aspects of our theoretical understanding of global scale magnetosphere-ionosphere coupling.

While the above general conclusions are concerned with the scheme of magnetosphere-ionosphere coupling as an integrated whole, we have also identified the following pressing issues which address specific areas of MIC:

- Global imagery of the aurora in optical, UV and X-ray is needed to determine a coherent picture of the global configurations of MIC.
- What is the specific spatial and temporal relationship between global-scale field-aligned current ( $J_{\parallel}$ ) systems, auroral potential structures and arc structures?
- What and where are the generators of the entire scale spectrum of field-aligned currents? In particular, is the MIC dynamo a voltage generator or a current generator? What is the nature of the auroral return current? How does upper atmospheric circulation affect magnetospheric current flow?
- Because auroral phenomena are generally associated with topological boundaries of magnetospheric regions, three-dimensional models of the magnetospheric electric field structure is needed.
- Quantitative models of ionospheric influence upon magnetospheric phenomena are needed. In particular, the importance of the ionospheric plasma source and of the influence of ionospheric conductivity upon MIC dynamics must be addressed quantitatively.
- At present there is a paucity of theories of auroral morphology in relation to MIC processes in the dayside and in the post-midnight sectors.
- It is by now clear that auroral plasma kinetic characteristics show both adiabatic and diffusive signatures. However, the feedback effects of wave turbulence generated by the auroral electron beam upon the distribution of  $E_{\parallel}$  have not been quantitatively clarified.
- Currently, there is a noticeable lack of theories of dynamic auroral processes such as the westward traveling surge.
- Since auroral arcs at very high latitude regions can involve plasmas from the magnetosheath and from the plasma sheet, quantitative theories of formation for such arcs are needed to understand their relationship to the oval arcs.
- Auroral kilometric radiation (AKR) reflects detailed structures of plasma processes of its source region, therefore, AKR can be used as a remote diagnostic tool of MIC processes. Once terrestrial AKR processes are understood, this diagnostic tool can be applied to MIC processes in other planets.

- Auroral wave-particle interactions depend on kinetic characteristics of the plasma. Some of these characteristics may result from the plasma transport itself, therefore, it is important to quantitatively understand the kinetic theory of auroral plasma transport.

## VI. REFERENCES

- Akasofu, S. -I., Polar and magnetospheric substorms, p. 223, D. Reidel, Dordrecht, Netherlands, 1968.
- Akasofu, S. -I., and J. R. Kan, Dayside and nightside auroral arc systems Geophys. Res. Lett., **7**, 753, 1980. Akasofu, S.-I., Recent progress in studies of DMSP auroral photographs, Space Sci. Rev., **19**, 169, 1976.
- Akasofu, S. -I., and J. R. Kan, Dayside and nightside auroral arc systems, Geophys. Res. Lett., **7**, 753, 1980.
- Anderson, H. L., Birkeland currents and auroral structure, J. Geomag. Geoelec., **30**, 381, 1978.
- Anderson, R. R., ISEE observations of second harmonic auroral kilometric radiation, J. Geophys. Res., submitted, 1983.
- Arnoldy, R. L., Review of auroral particle precipitation, Physics of Auroral Arc Formation, ed. by S. -I. Akasofu and J. R. Kan, p. 56, Geophysical Monograph Series, Vol. 25, American Geophysical Union, Washington, D.C., 1981.
- Ashour-Abdalla, M., H. Okuda, and C. Z. Cheng, Acceleration of heavy ions on auroral field lines, Geophys. Res. Lett., **8**, 795, 1981.
- Axford, W. I., and C. O. Hines, A unifying theory of high-latitude geophysical phenomena and geomagnetic storms, Can. J. Phys., **39**, 1433, 1961.
- Banks, P. M., Magnetosphere, ionosphere and atmosphere interactions, in Solar Plasma Physics, Vol. II, ed. by C. F. Kennel, L. J. Lanzerotti and E. N. Parker, p. 57, North Holland, Amsterdam, 1979.
- Banks, P. M., and T. E. Holzer, High-latitude plasma transport: the polar wind, J. Geophys. Res., **75**, 6, 1969.
- Barbosa, D. D., Electrostatic mode coupling at  $2 \omega_{UH}$ : A generation mechanism for auroral kilometric radiation, Ph.D. dissertation, Dept. of Phys. Univ. of Calif., Los Angeles, 1976.
- Benson, R. F., Source mechanism for terrestrial kilometric radiation, Geophys. Res. Lett., **2**, 52, 1975.
- Benson, R. F., Auroral kilometric radiation source region observations from ISIS 1, Physics of Auroral Arc Formation, ed. by S.-I Akasofu and J. R. Kan, p. 369, Geophysical Monograph Series, Vol. 25, American Geophysical Union, 1981.

- Benson, R. F., Harmonic auroral kilometric radiation of natural origin, Geophys. Res. Lett., **9**, 1120, 1982.
- Benson, R. F., and W. Calvert, ISIS-1 observations at the source of auroral kilometric radiation, Geophys. Res. Lett., **6**, 479, 1979.
- Benson, R. F., W. Calvert, and D. M. Klumpar, Simultaneous wave and particle observations in the auroral kilometric radiation source region, Geophys. Res. Lett., **7**, 959, 1980.
- Bernstein, W., B. Hultqvist, and H. Borg, Some implications of low altitude observations of isotropic precipitation of ring current protons beyond the plasma pause, Planet. Space Sci., **22**, 767, 1974.
- Bunn, F. E., and G. G. Shepherd, The solar terrestrial event of 14-21 December 1971: The pattern of 6300Å emission over the polar cap, Planet. Space Sci., **27**, 973, 1979.
- Burke, W. J., M. S. Gussenhoven, M. C. Kelley, D. A. Hardy, and F. J. Rich, Electric and magnetic field characteristics of discrete arcs in the polar cap, J. Geophys. Res., **87**, 2431, 1982.
- Burke, W. J., M. Silevitch, and D. A. Hardy, Observation of small-scale auroral vortices by S3-2 satellite, J. Geophys. Phys., **88**, 3127, 1983.
- Calvert, W., The signature of auroral kilometric radiation on ISIS-1 ionograms, J. Geophys. Res., **86**, 76, 1981a.
- Calvert, W., The AKR emission cone at low frequencies, Geophys. Res. Lett., **8**, 1159, 1981b.
- Calvert, W., A feedback model for the source of auroral kilometric radiation, J. Geophys. Res., **87**, 8199, 1982.
- Cattell, C. A., and M. K. Hudson, Flute mode waves near  $\omega_{lh}$  excited by ion rings in velocity space, Geophys. Res., **9**, 1167, 1982.
- Chang, T., and B. Coppi, Lower hybrid acceleration and ion evolution in the superauroral region, Geophys. Res. Lett., **8**, 1253, 1981.
- Chappell, C. R., C. R. Baugher, and J. L. Horwitz, New advances in thermal plasma research, Rev. Geophys. Space Phys., **18**, 853, 1980.
- Chappell, C. R., J. L. Green, J. F. E. Johnson, and J. H. Waite, Jr, Pitch angle variations in magnetospheric thermal plasma-initial observations from Dynamics Explorer-1, Geophys. Res. Lett., **9**, 933, 1982a.
- Chappell, C. R., R. C. Olsen, J. L. Green, J. F. E. Johnson, and J. H. Waite, Jr., The discovery of nitrogen ions in the Earth's magnetosphere, Geophys. Res. Lett., **9**, 937, 1982b.
- Chiu, Y. T., and J. M. Cornwall, Electrostatic model of a quiet auroral arc, J. Geophys. Res., **85**, 543, 1980.
- Chiu, Y. T., and D. J. Gorney, Eddy intrusion of hot plasma into the polar cap and formation of polar-cap arcs, Geophys. Res. Lett., **10**, 463, 1983.

- Chiu, Y. T., and M. Schulz, Self-consistent particle and parallel electrostatic field distributions in the magnetospheric-ionospheric auroral region, J. Geophys. Res., **83**, 629, 1978.
- Chiu, Y. T., A. L. Newman, and J. M. Cornwall, On the structures and mappings of auroral electrostatic potentials, J. Geophys. Res., **86**, 10029, 1981a.
- Chiu, Y. T., J. M. Cornwall, and M. Schultz, Effects of auroral-particle anisotropies and mirror forces on high-latitude electric fields, in Physics of Auroral Arc Formation, p. 234, ed. by S. -I. Akasofu and J. R. Kan, Geophysical Monograph Series, Vol. 25, American Geophysical Union, 1981, Washington, D. C., 1981b.
- Chiu, Y. T., J. M. Cornwall, J. F. Fennell, D. J. Gorney, and P. F. Mizera, Auroral plasmas in the evening sector: Satellite observations and theoretical interpretations, Space Sci. Rev., **35**, 211, 1983a.
- Chiu, Y. T., M. Schulz, J. F. Fennell, and A. M. Kishi, Mirror instability and the origin of morningside auroral structure, J. Geophys. Res., **88**, 4041, 1983b.
- Collin, H. L., R. D. Sharp, E. G. Shelly, and R. G. Johnson, Some general characteristics of upflowing ion beams over the auroral zone and their relationship to auroral electrons, J. Geophys. Res., **86**, 6820, 1981.
- Cornwall, J. M., and Y. T. Chiu, Ion distribution effects of turbulence on a kinetic auroral arc model, J. Geophys. Res., **87**, 1517, 1982.
- Cowley, S. W. H., Plasma populations in a simple open model magnetosphere, Space Sci. Reviews, **26**, 217, 1980.
- Cowley, S. W. H., and M. Ashour-Abdalla, Adiabatic plasma convection in a dipole field: Electron forbidden-zone effects for a simple electric model, Planet. Space Sci., **24**, 805-819, 1976a.
- Cowley, S. W. H., and M. Ashour-Abdalla, Adiabatic plasma convection in a dipole field: Proton forbidden-zone effects for a simple electric model, Planet. Space Sci., **24**, 821-833, 1976b.
- Davis, T. N., The morphology of the auroral displays of 1957-1958, 2. Detailed analysis of Alaska data and analysis of high-latitude data, J. Geophys. Res., **67**, 75, 1962.
- DeCoster, R. J., and L. A. Frank, Observations pertaining to the dynamics of the plasma sheet, J. Geophys. Res., **84**, 5099, 1979.
- Dungey, J. W., Interplanetary magnetic field and the auroral zone, Phys. Rev. Lett., **6**, 47, 1961.
- Dusenberry, P. B., and L. R. Lyons, Generation of ion-conic distribution by upgoing ionospheric electrons, J. Geophys. Res., **86**, 7627, 1981.
- Dusenberry, P. B., and L. R. Lyons, General concepts on the generation of auroral kilometric radiation, J. Geophys. Res., **87**, 7467, 1982.
- Ejiri, M., Trajectory traces of charged particles in magnetosphere, J. Geophys. Res., **85**, 4798, 1978.

- Ejiri, M., R. A. Hoffman, and P. H. Smith, Energetic particle penetration into the inner magnetosphere, J. Geophys. Res., **85**, 653, 1980.
- Evans, D. S., Precipitating electron fluxes formed by a magnetic field-aligned potential difference, J. Geophys. Res., **79**, 2853, 1974.
- Evans, D. S., The acceleration of charged particles at low altitudes, in Physics of Solar Planetary Environments, ed. D. J. Williams, Amer. Geophysics Union, Washington, D. C., 730, 1976.
- Evans, D. S., N. C. Maynard, J. Troim, T. Jacobsen, and A. Egeland, Auroral vector electric field and particle comparisons 2. Electrodynamics of an arc, J. Geophys. Res., **82**, 2235, 1977.
- Evans, D. S., and T. E. Moore, Precipitating electrons associated with the diffuse aurora: Evidence for electrons of atmospheric origin in the plasma sheet, J. Geophys. Res., **84**, 6451, 1979.
- Fejer, J. A., Theory of geomagnetic daily disturbance variations, J. Geophys. Res., **69**, 123, 1964.
- Fennell, J. F., P. M. Mizera, and D. R. Croley, Jr., Observations of ion and electron distributions during the July 29 and July 30 storm period, in Proceedings of Magnetospheric Boundary Layers Conference, Alpach, ESA/SP-148, 1979.
- Fennell, J. F., D. J. Gorney, and P. F. Mizera, Auroral particle distribution functions and their relationship to inverted V's and auroral arcs, in Physics of Auroral Arc Formation, ed. by S.-I. Akasofu and J. R. Kan, Geophysical Monograph Series, Vol. 25, p. 91, American Geophysical Union, Washington, D. C., 1981.
- Frank, L. A., and K. L. Ackerson, Observation of charged particle precipitation into the auroral zone, J. Geophys. Res., **76**, 3612, 1971.
- Frank, L. A., and D. A. Gurnett, Distributions of plasmas and electric fields over the auroral zones and polar caps, J. Geophys. Res., **76**, 6829, 1971.
- Frank, L. A., J. D. Craven, J. L. Burch, and J. D. Winningham, Polar views of the earth's aurora with Dynamics Explorer, Geophys. Res. Lett., **9**, 1001, 1982.
- Freund, H. P., I. Haber, P. Palmadesso, and K. Papadopoulos, Strongly turbulent stabilization of electron beam-plasma interactions, Phys. Fluids, **23**, 518, 1980.
- Fridman, M., and J. Lemaire, Relationship between auroral electron fluxes and field-aligned electric potential difference, J. Geophys. Res., **85**, 664, 1980.
- Geiss, J., H. Balsiger, P. Eberhardt, H. P. Walker, L. Weber, and D. T. Young, Dynamics of magnetosphere ion composition as observed by the GEOS mass spectrometer, Space Sci. Rev., **22**, 537, 1978.
- Ghielmetti, A. G., R. G. Johnson, R. D. Sharp, and E. G. Shelley, The latitudinal, diurnal and altitudinal distributions of upward flowing energetic ions of ionospheric origin, Geophys. Res. Lett., **5**, 59, 1978.
- Ghielmetti, A. G., R. D. Sharp, E. G. Shelley, and R. G. Johnson, Downward flowing ions and evidence for injection of ionospheric ions into the plasma sheet, J. Geophys. Res., **84**, 5781, 1979.

- Gonzales, C., M. Kelley, B. Fejer, J. Vickrey, and R. Woodman, Equatorial electric fields during magnetically disturbed conditions, 2, Implications of simultaneous auroral and equatorial measurements, J. Geophys. Res., **84**, 5803, 1979.
- Gorney, D. J., A. Clarke, D. R. Croley, J. F. Fennell, J. G. Luhmann, and P. F. Mizera, The distribution of ion beams and conics below 8000 km, J. Geophys. Res., **86**, 83, 1981.
- Gorney, D. J., S. R. Church, and P. F. Mizera, On ion harmonic frequency structure in auroral zone waves: The effect of ion conic damping of auroral hiss, J. Geophys. Res., **87**, 10479, 1982.
- Grabbe, C. L., Theory of the fine structure of auroral kilometric radiation, Geophys. Res. Lett., **9**, 155, 1982.
- Grabbe, C. L., K. Papadopoulos, and P. J. Palmadesso, A coherent nonlinear theory of auroral kilometric radiation, 1. Steady state model, J. Geophys. Res., **85**, 3337, 1980.
- Green, J. L., Observations pertaining to the generation of auroral kilometric radiation, in Physics of Auroral Arc Formation, ed. by S.-I. Akasofu and J. R. Kan, p. 359, Geophysical Monograph Series, Vol. 25, American Geophysical Union, Washington, D. C., 1981.
- Green, J. L., D. A. Gurnett, and R. A. Hoffman, A correlation between auroral kilometric radiation and inverted-V electron precipitation, J. Geophys. Res., **84**, 5216, 1979.
- Green, J. L., N. A. Saflekos, D. A. Gurnett, and T. A. Potemra, A correlation between auroral kilometric radiation and field-aligned currents, J. Geophys. Res., **87**, 10463, 1982.
- Greenwald, R. A., Recent advances in magnetosphere-ionosphere coupling, Rev. Geophys. Space Phys., **20**, 577, 1982.
- Gurgiolo, C., and J. L. Burch, DE-1 observations of the polar wind - A heated and an unheated component, Geophys. Res. Lett., **9**, 945, 1982.
- Gurnett, D. A., The Earth as a radio source: Terrestrial kilometric radiation, J. Geophys. Res., **79**, 4227, 1974.
- Gurnett, D. A., Electromagnetic plasma wave emission from the auroral field lines, in Auroral Processes, ed. by C. T. Russell, p. 91, Center for Academic Publications, Japan, Japanese Scientific Societies Press, Tokyo, 1979.
- Gurnett, D. A., and R. R. Anderson, The kilometric radio emission spectrum: Relationship to auroral acceleration processes, in Physics of Auroral Arc Formation, ed. by S.-I. Akasofu and J. R. Kan, p. 341, Geophysical Monograph Series, Vol. 25, American Geophysical Union, 1981.
- Gurnett, D. A., and L. A. Frank, Observed relationships between electric fields and auroral particle precipitation, J. Geophys. Res., **78**, 145, 1973.
- Gurnett, D. A., and L. A. Frank, A region of intense plasma wave turbulence on auroral field lines, J. Geophys. Res., **82**, 1031, 1977.



- Gurnett, D. A., and J. L. Green, On the polarization and origin of auroral kilometric radiation, J. Geophys. Res., **83**, 689-696, 1978.
- Gurnett, D. A., R. R. Anderson, F. L. Scarf, R. W. Fredricks, and E. J. Smith, Initial results from the ISEE-1 and -2 plasma wave investigation, Space Sci. Reviews, **23**, 103, 1979.
- Haerendel, G., and G. Paschmann, Interaction of the solar wind with the dayside magnetosphere, in Magnetosphere Plasma Physics, ed. by A. Nishida, p. 49, D. Reidel, Dordrecht, Holland, 1982.
- Hasegawa, A., and T. Sato, Generation of field-aligned current during substorms, in Dynamics of the Magnetosphere, ed. by S.-I. Akasofu, p. 259, D. Reidel, Dordrecht, Holland, 1979.
- Heelis, R. A., and W. B. Hanson, High-latitude ion convection in the nighttime F-region, J. Geophys. Res., **85**, 1995, 1980.
- Heppner, J. P., M. L. Miller, M. B. Pongratz, G. M. Smith, L. L. Smith, S. B. Mende, and N. R. Nath, The Cameo barium releases: E-fields over the polar cap, J. Geophys. Res., **86**, 3519, 1981.
- Hoffman, R. A., and C. S. Lin, Study of inverted-V auroral precipitation events, in Physics of Auroral Arc Formation, pp. 80-90, ed. S.-I. Akasofu and J. R. Kan, Geophysical Monograph Series, Vol. 25, American Geophysical Union, Washington, D. C., 1981.
- Horwitz, J. L., The ionosphere as a source of magnetospheric ions, Rev. Geophys. Space Phys., **20**, 929, 1982.
- Horwitz, J. L., and S.-I. Akasofu, On the relationship of the polar cap current system to the north-south component of the interplanetary magnetic field, J. Geophys. Res., **84**, 2567, 1979.
- Horwitz, J. L., C. R. Baugher, C. R. Chappell, E. G. Shelley, and D. T. Young, Conical pitch angle distributions of very low-energy ion fluxes observed by ISEE-1., J. Geophys. Res., **87**, 2311, 1982.
- Hudson, M. K., and F. S. Mozer, Electric shocks, double layers and anomalous resistivity in the magnetosphere, Geophys. Res. Lett., **5**, 131, 1978.
- Hultqvist, B., Recent progress in understanding of the ion composition in the magnetosphere and some major question marks, Rev. Geophys. Sp. Phys., **20**, 589, 1982.
- Iijima, T., and T. A. Potemra, Large-scale characteristics of field-aligned currents associated with substorms, J. Geophys. Res., **83**, 599, 1978.
- Ismail, S., and C.-I. Meng, A classification of polar cap auroral arcs, Planet. Space Sci., **30**, 319, 1982.
- Johnson, R. G., Energetic ion composition in the Earth's magnetosphere, Rev. Geophys. Space Physics, **17**, 696, 1979.
- Johnstone, A. D., Pulsating auroras, Nature, **274**, 119, 1978.

- Kaiser, M. L., J. K. Alexander, A. C. Riddle, J. B. Pearce, and J. W. Warwick, Direct measurements of the polarization of terrestrial kilometric radiation from Voyagers 1 and 2, Geophys. Res. Lett., **5**, 857, 1978.
- Kamide, Y., and G. Rostoker, The spatial relationships of field-aligned currents and auroral electrojets to the distribution of nightside aurora, J. Geophys. Res., **82**, 5589, 1977.
- Kamide, Y., and S. Matsushita, Simulation studies of ionospheric electric fields and currents in relation to field-aligned currents, 2, Substorms, J. Geophys. Res., **84**, 4099, 1979.
- Kamide, Y., A. D. Richmond, and S. Matsushita, Estimation of ionospheric electric fields, ionospheric currents, and field-aligned currents from ground magnetic records, J. Geophys. Res., **86**, 801, 1981.
- Karney, C. F. F., Stochastic ion heating by a lower hybrid wave, Phys. Fluids., **21**, 1584, 1978.
- Kaufman, R. L., and P. M. Kintner, Upgoing ion beams, 1. Microscopic analysis, J. Geophys. Res., **87**, 487, 1982.
- Kaye, S. M., and M. G. Kivelson, Time-dependent convection electric fields and plasma injection, J. Geophys. Res., **84**, 4183, 1979.
- Kaye, S. M., R. G. Johnson, R. D. Sharp, and E. G. Shelley, Observations of transient  $H^+$  and  $O^+$  bursts in the equatorial magnetosphere, J. Geophys. Res., **86**, 1335, 1981.
- Kennel, C. F., and M. Ashour-Abdalla, Electrostatic waves and the strong diffusion of magnetospheric electrons, in Magnetospheric Plasma Physics, edited by A. Nishida, D. Reidel, Boston, Ma., 1982.
- Kennel, C. F., F. L. Scarf, R. W. Fredricks, J. H. McGehee, and F. V. Coronitti, ULF electric field observations in the magnetosphere, J. Geophys. Res., **75**, 6136, 1970.
- Kindel, J. M., and C. F. Kennel, Topside current instabilities, J. Geophys. Res., **76**, 3055, 1971.
- Kintner, P. M., and D. J. Gorney, A search for the plasma processes associated with the source of ion conics, J. Geophys. Res., in press, 1983.
- Kintner, P. M., M. C. Kelley, and F. S. Mozer, Electronics hydrogen cyclotron waves near one Earth radius altitude in the polar magnetosphere, Geophys. Res. Lett., **51**, 139, 1978.
- Kivelson, M. G., and D. J. Southwood, Local time variations of particle flux produced by an electrostatic field in the magnetosphere, J. Geophys. Res., **80**, 56-65, 1975.
- Kleitzing, C., C. Cattell, F. S. Mozer, S. -I. Akasofu and K. Makita, Evidence for electrostatic shocks as a source of discrete auroral arcs, J. Geophys. Res., **88**, 4105, 1983.
- Klumpar, D., Relationships between auroral particle distributions and magnetic field perturbations associated with field-aligned currents, J. Geophys. Res., **84**, 6524, 1979a.

- Klumpar, D. M., Transversely accelerated ions: An ionospheric source of hot magnetospheric ions, J. Geophys. Res., **84**, 7201, 1979b.
- Knight, S., Parallel electric fields, Planet. Space Sci., **21**, 741, 1973.
- Koons, H. C., and J. F. Fennell, Particle and wave dynamics during plasma injections, J. Geophys. Res., in press, 1983.
- Kurth, W. S., L. A. Frank, M. Ashour-Abdalla, D. A. Gurnett, and G. G. Burek, Observations of a free-energy source for intense electrostatic waves, Geophys. Res. Lett., **7**, 293, 1980.
- Lemaire, J., and M. Scherer, Ionosphere-plasma sheet field aligned currents and parallel electric fields, Planet. Space Sci., **22**, 1485, 1974.
- Lennartsson, W., On high latitude convection field inhomogeneities; parallel electric fields and inverted-V precipitation events, Planet. Space Sci., **25**, 89, 1977.
- Lennartsson, W., On the consequences of the interaction between the auroral plasma and the geomagnetic field, Planet. Space Sci., **28**, 135, 1980.
- Lunblad, J. A., F. Soraas, and K. Aarsnes, Substorm morphology of >100 keV protons, Planet Space Sci., **27**, 841, 1979.
- Lundin, R., B. Hultqvist, E. Dubinin, A. Zackarov, and N. Pissarenko, Observations of outflowing ion beams on auroral field lines at altitudes of many earth radii, Planet. Space Sci., **30**, 715, 1982.
- Lyons, L. R., Electron diffusion driven by magnetospheric electrostatic waves, J. Geophys. Res., **79**, 575, 1974.
- Lyons, L. R., Generation of large-scale regions of auroral currents, electric potentials, and precipitation by the divergence of the convection electric field, J. Geophys. Res., **86**, 17, 1980.
- Lyons, L. R., Discrete aurora as the direct result of an inferred, high-altitude generating potential distribution, J. Geophys. Res., **86**, 1, 1981a.
- Lyons, L. R., The field-aligned current versus electric potential relation and auroral electrodynamics, in Physics of Auroral Arc Formation, ed. by S.-I. Akasofu and J. R. Kan, American Geophys. Union, Washington, D. C., 252, 1981b.
- Lyons, L. R., and T. W. Speiser, Evidence for current-sheet acceleration in the geomagnetic tail, J. Geophys. Res., **87**, 2276, 1982.
- Lyons, L. R., D. S. Evans, and R. Lundin, An observed relation between magnetic field-aligned electric fields and downward electron energy fluxes in the vicinity of auroral forms, J. Geophys. Res., **84**, 457, 1979.
- Lysak, R. L. and C. W. Carlson, Effect of microscopic turbulence on magnetosphere-ionosphere coupling, Geophys. Res. Lett., **8**, 269, 1981.

- Lysak, R. L., and C. T. Dum, Dynamics of magnetosphere-ionosphere coupling including turbulent transport, J. Geophys. Res., **88**, 365, 1983.
- Lysak, R. L., M. K. Hudson, and M. Temerin, Ion heating by strong electrostatic ion cyclotron turbulence, J. Geophys. Res., **85**, 678, 1980.
- Mawson, D., Records of the aurora polaris, Australian Antarctic Expedition, 1911-1914, Sci. Rep., Ser. B, p. 11, 1925.
- Maynard, N. C., D. S. Evans, B. Maehlum, and A. Egeland, Auroral vector electric field and particle comparisons, 1. Premidnight convection topology, J. Geophys. Res., **82**, 2227, 1977.
- Maynard, N., T. Aggson, and J. Heppner, Magnetosphere observation of large sub-auroral electric fields, Geophys. Res. Lett., **7**, 881-884, 1980.
- Melrose, D. B., An interpretation of Jupiter's radiation and the terrestrial kilometric radiation as direct amplified gyroemission, Astro. Phys. J., **207**, 651, 1976.
- Melrose, D. B., K. G. Ronnmark, and R. G. Hewitt, Terrestrial kilometric radiation: The cyclotron theory, J. Geophys. Res., in press, 1982.
- Meng, C. -I., and S. -I. Akasofu, The relation between the polar cap auroral arc and the auroral oval arc, J. Geophys. Res., **81**, 4004, 1976.
- Mizera, P. F., and J. F. Fennell, Signatures of electric fields from high and low altitude particle distributions, Geophys. Res. Lett., **4**, 311, 1977.
- Mizera, P. F., J. F. Fennell, D. R. Croley, Jr., A. Vampola, F. S. Mozer, R. B. Torbert, M. Temerin, R. L. Lysak, M. K. Hudson, C. A. Cattell, R. G. Johnson, R. D. Sharp, A. Ghielmetti, and P. M. Kintner, Aurora inferred from S3-3 particles and fields, J. Geophys. Res., **86**, 2329, 1981.
- Moore, T. E., Modulation of terrestrial ion escape flux composition (by low-altitude acceleration and charge exchange chemistry), J. Geophys. Res., **85**, 2011, 1980.
- Mozer, F. S., C. W. Carlson, M. K. Hudson, R. B. Torbert, B. Parady, J. Yatteau, and M. C. Kelley, Observations of paired electrostatic shocks in the polar magnetosphere, Phys. Rev. Lett., **38**, 292, 1977.
- Mozer, F. S., C. A. Cattell, M. K. Hudson, R. L. Lysak, M. Temerin, and R. B. Torbert, Satellite measurements and theories of low-altitude auroral particle acceleration, Space Sci. Reviews, **27**, 155, 1980.
- Nisbet, J. S., M. J. Miller, and L. A. Carpenter, Currents and electric fields in the ionosphere due to field-aligned auroral currents, J. Geophys. Res., **83**, 2647, 1978.
- Oguti, T., Television observations of auroral arcs, in Physics of Auroral Arc Formation, ed. S. -I. Akasofu and J. R. Kan, p. 31-41, AGU Monograph 25, AGU, Washington, D. C., 1981.
- Okuda, H., and M. Ashour-Abdalla, Formation of a conical distribution and intense ion heating in the presence of hydrogen cyclotron waves, Geophys. Res. Lett., **8**, 811, 1981.

- Omidi, N., and D. A. Gurnett, Growth rate calculations of auroral kilometric radiation using the relativistic resonance condition, J. Geophys. Res., **87**, 2377, 1982.
- Oya, H., and A. Morioka, Observational evidence of Z-mode waves as the origin of auroral kilometric radiation based on AKR data detected by JIKIKEN (EXOS-B) satellite, J. Geophys. Res., submitted, 1982.
- Palmadesso, P. J., T. P. Coffey, S. L. Ossakow, and K. Papadopoulos, Topside ionospheric ion heating due to electrostatic ion cyclotron turbulence, Geophys. Res. Lett., **3**, 105, 1974.
- Papadopoulos, K., and P. Palmadesso, Excitation of lower hybrid waves in a plasma by electron beams, Phys. Fluids, **19**, 605, 1976.
- Papadopoulos, K., J. D. Gaffey, Jr., and P. J. Palmadesso, Stochastic acceleration of large M/Q ions by hydrogen cyclotron waves in the magnetosphere, Geophys. Res. Lett., **7**, 1014, 1980.
- Peterson, W. K., R. D. Sharp, E. G. Shelley, R. G. Johnson, and H. Balsiger, Energetic ion composition of the plasma sheet, J. Geophys. Res., **86**, 761, 1981.
- Potemra, T. A., Current systems in the Earth's magnetosphere, IUGG Quadrennial Report, Rev. Geophys. Space Phys., **17**, 640, 1979.
- Potemra, T. A., Magnetospheric Currents, Geophysical Monograph Series, Vol. 25, American Geophysical Union, Washington, D. C., in press, 1983.
- Rairden, R. L., L. A. Frank, and J. D. Craven, Geocoronal imaging with Dynamics Explorer: A first look, J. Geophys. Res., accepted, 1983.
- Richardson, J. D., J. F. Fennell, and D. R. Croley, Jr., Observations of field-aligned ion and electron beams from SCATHA (P78-2), J. Geophys. Res., **86**, 10015, 1981.
- Ronnmark, K., H. Borg, P. H. Christensen, M. P. Gough, and D. Jones, Banded electron cyclotron harmonic instability - A first comparison of theory and experiment, Space Sci. Reviews, **22**, 401, 1978.
- Rostoker, G., and R. Bostrom, A mechanism for driving the gross Birkeland current configuration in the auroral oval, J. Geophys. Res., **81**, 235, 1976.
- Roth, I., and M. K. Hudson, Particle simulations of electrostatic emissions near the lower hybrid frequency, J. Geophys. Res., **88**, 483, 1983.
- Roux, A., and R. Pellat, Coherent generation of the auroral kilometric radiation by nonlinear beatings between electrostatic waves, J. Geophys. Res., **84**, 5189, 1979.
- Rowland, H. L., and P. J. Palmadesso, Anomalous resistivity due to low frequency turbulence, submitted to J. Geophys. Res., 1983.
- Rowland, H. L., P. J. Palmadesso, and K. Papadopoulos, Anomalous resistivity on auroral field lines, Geophys. Res. Lett., **8**, 1257, 1981.
- Royrvik, O., and T. N. Davis, Pulsating auroras: Local and global morphology, J. Geophys. Res., **82**, 4720, 1977.

- Saflekos, N. A., R. E. Sheehan, and R. L. Carovillano, Global nature of field-aligned currents and their relation to auroral phenomena, Rev. Geophys. Space Phys., **20**, 709, 1982.
- Schild, M. A., J. W. Freeman, and A. J. Dessler, A source for field-aligned currents at auroral latitudes, J. Geophys. Res., **74**, 247, 1969.
- Sharber, J. R., The continuous (diffuse) aurora and auroral-E ionization, in Physics of Space Plasmas, ed. by T. S. Chang, B. Coppi, and J. R. Jasperse, Scientific Publishers, Cambridge, Ma., in press, 1981 .
- Sharp, R. D., E. G. Shelley, R. G. Johnson, and A. G. Ghielmetti, Counterstreaming electron beams at altitudes of  $\sim 1$  Re over the auroral zone, J. Geophys. Res., **85**, 92, 1980.
- Sharp, R. D., R. G. Johnson, and E. G. Shelley, Observation of an ionospheric acceleration mechanism producing energetic (keV) ions primarily normal to the geomagnetic field direction, J. Geophys. Res., **82**, 3324, 1977.
- Sharp, R. D., W. L. Lennartsson, W. K. Peterson, and E. Ungstrup, The mass dependence of wave-particle interactions as observed with the ISEE-1 energetic ion spectrometer, Geophys. Res. Lett., in press, 1983.
- Shawhan, S. D., DE-1 auroral kilometric radiation observations, J. Geophys. Res., submitted, 1983.
- Shawhan, S. D., and D. A. Gurnett, Polarization measurements of auroral kilometric radiation by Dynamics Explorer-1, Geophys. Res. Lett., **9**, 913, 1982.
- Shelley, E. G., R. G. Johnson, and R. D. Sharp, Satellite observations of energetic heavy ions during a geomagnetic storm, J. Geophys. Res., **77**, 6104, 1972.
- Shelley, E. G., R. D. Sharp, and R. G. Johnson, Satellite observations of an ionospheric acceleration mechanism, Geophys. Res. Lett., **6**, 54, 1976.
- Shelley, E. G., W. K. Peterson, A. Ghielmetti, and J. Geiss, The polar ionosphere as a source of energetic magnetospheric plasma, Geophys. Res. Lett., **9**, 941, 1982.
- Singh, N., R. W. Schunk, and J. J. Sojka, Energization of ionospheric ions by electrostatic hydrogen cyclotron waves, Geophys. Res. Lett., **8**, 1249, 1981.
- Siscoe, G. L., Energy coupling between regions 1 and 2 Birkeland current systems, J. Geophys. Res., **87**, 5124, 1982.
- Sonnerup, B. U. Ö., Theory of the low-latitude boundary layer, J. Geophys. Res., **85**, 2017, 1980.
- Speiser, T. W., Particle trajectories in model current sheets, 1. Analytical solutions, J. Geophys. Res., **70**, 4219, 1965.
- Stern, D. P., Large-scale electric fields in the Earth's magnetosphere, Rev. Geophys. Space Phys., **15**, 156, 1977.
- Swift, D. W., The effect of the neutral sheet on magnetospheric plasma, J. Geophys. Res., **82**, 1288, 1977.

- Sugira, M., N. C. Maynard, W. H. Farthing, J. P. Heppner, B. G. Ledley, and L. J. Cahill, Jr., Initial results on the correlation between the magnetic and electric fields observed from the DE-2 satellite in the field-aligned current region, Geophys. Res. Lett., **9**, 985, 1983.
- Temerin, M., The polarization, frequency, and wavelengths of high-latitude turbulence, J. Geophys. Res., **83**, 2609, 1978.
- Temerin, M., Plasma waves on auroral field lines, Physics of Auroral Arc Formation, ed. by S. -I. Akasofu and J. R. Kan, p. 351, Geophysical Monograph Series, Vol. 25, American Geophysical Union, Washington, D. C., 1981.
- Temerin, M., C. Cattell, R. Lysak, M. Hudson, R. B. Torbert, F. S. Mozer, R. D. Sharp, and P. M. Kintner, The small-scale structure of electrostatic shocks, J. Geophys. Res., **86**, 11278, 1981.
- Temerin, M., K. Cerny, W. Lotko, and F. S. Mozer, Observations of double layers and solitary waves in the auroral plasma, Phys. Rev. Lett., **48**, 1175, 1982.
- Torbert, R. B., and F. S. Mozer, Electrostatic shocks as the source of discrete arcs, Geophys. Res. Lett., **5**, 135, 1978.
- Ungstrup, E., D. M. Klumpar, and W. J. Heikkila, Heating of ions to superthermal energies in the topside ionosphere by electrostatic ion cyclotron waves, J. Geophys. Res., **84**, 4289, 1979.
- Vasyliunas, V. M., Mathematical models of magnetospheric convection and its coupling to the ionosphere, in Particles and Fields in the Magnetosphere, ed. by B. McCormac, p. 60, D. Reidel, Dordrecht-Holland, 1970.
- Vondrak, R., and F. Rich, Simultaneous Chatanika radar and S3-2 satellite measurements of ionospheric electrodynamics in the diffuse aurora, J. Geophys. Res., **87**, 6173, 1982.
- Whalen, B. A., and I. B. McDiarmid, Pitch angle diffusion of low-energy auroral electrons, J. Geophys. Res., **78**, 1608, 1973.
- Whalen, B. A., W. Bernstein, and D. W. Daly, Low altitude acceleration of ionospheric ions, Geophys. Res. Lett., **5**, 55, 1978.
- Williams, D. J., The Earth's ring current: Causes, generation, and decay, Space Sci. Reviews, in press, 1983.
- Winningham, J. D., and C. Gurgiolo, DE-2 photoelectron measurements consistent with a large scale parallel electric field over the polar cap, Geophys. Res. Lett., **9**, 977, 1982.
- Wolf, R. A., Ionosphere-magnetosphere coupling, Space Sci. Rev., **17**, 537, 1975.
- Wolf, R. A., M. Harel, R. W. Spiro, G. H. Voigt, P. H. Reiff, and C. K. Chen, Computer simulation of inner magnetospheric dynamics for the magnetic storm of July 29, 1977, J. Geophys. Res., **87**, 5949, 1982.
- Wu, C. S., and L. C. Lee, A theory of the terrestrial kilometric radiation, Astrophys. J., **230**, 621, 1979.

- Yang, W. H., and J. R. Kan, Generation of conic ions by auroral electric fields, J. Geophys. Res., **88**, 465, 1983.
- Young, D. T., H. Balsiger, and J. Geiss, Observed increase in the abundance of kilovolt  $O^+$  in the magnetosphere due to solar cycle effects, Adv. Space Res., **1**, 309, 1981.
- Young, T. S. T., J. D. Callen, and J. E. McCune, High-frequency electrostatic waves in the magnetosphere, J. Geophys. Res., **78**, 1082, 1973.



# CHAPTER 8

## SUBSTORMS IN THE MAGNETOSPHERE

### WORKING GROUP MEMBERS\*

D. N. Baker, Chairman  
*Los Alamos National Laboratory*

S. -I. Akasofu  
*University of Alaska*

W. Baumjohann  
*MPI fur Extraterrestrische Physik*

J. W. Bieber  
*Bartol Research Foundation*

D. H. Fairfield  
*NASA/Goddard Space Flight Center*

E. W. Hones, Jr.  
*Los Alamos National Laboratory*

B. Mauk  
*Johns Hopkins University*

R. L. McPherron  
*University of California, Los Angeles*

T. E. Moore  
*University of New Hampshire*

---

\*The members of Working Group 8 gratefully acknowledge the advice and written input of M. Abdalla and T. Terasawa who were unable to participate in the Workshop deliberations.

## CHAPTER 8

### SUBSTORMS IN THE MAGNETOSPHERE

I.	Introduction	8-3
II.	Solar Wind - Magnetosphere Coupling	8-7
III.	Driven Processes in Substorms	8-11
IV.	The Unloading Process in Substorms	8-18
V.	Summary and Future Directions	8-42
VI.	References	8-45

## I. INTRODUCTION

All large-scale objects orbiting the Sun have magnetospheres. The magnetospheres of those planets having substantial intrinsic magnetic fields (e.g., Mercury, Earth, Jupiter, Saturn) have been called "true magnetospheres" (Vasyliunas, 1982) and are formed by tangential stresses that the solar wind applies as it gives up momentum to the outer regions of the intrinsic fields; in the momentum exchange a system of electric currents is set up whose  $\mathbf{j} \times \mathbf{B}$  force effectively distorts the outer fringe of the magnetic field pulling it downstream to form a comet-like magnetotail. Solar wind plasma becomes entrained in the tail and carries the current that maintains the tail-like magnetic structure. The magnetospheres of comets and of planets that lack a substantial intrinsic field (e.g., Venus) have been called "magnetosphere-like systems" (Vasyliunas, 1982) and are formed by the "hanging-up" of interplanetary magnetic field lines in the envelope of conducting ionized gas (ionosphere) created by solar ultra-violet radiation and impacting solar wind plasma. These temporarily restrained magnetic field lines are also pulled downstream to form a magnetotail.

The continual flow of the solar wind and its continual imparting of energy to the magnetospheres of these solar system objects requires that the energy be dissipated in some manner, continually and/or intermittently, possibly to the parent body or back to the solar wind, or to both. Evidence for this dissipation has been found in the Earth's magnetosphere where dissipation of energy in the ring current and the auroral ionosphere has been observed. Additionally, it appears that, intermittently, a portion of the plasma sheet (the plasma-containing region at the midplane of the tail) is severed from the Earth and flows away with the solar wind. The sudden occurrence of this magnetospheric dissipation process is taken as one basic underlying feature of the collection of processes that we call a magnetospheric substorm.

As a starting point in the discussion of substorms our committee strove for a mutually acceptable and (necessarily) broad definition of a substorm. At the beginning, there was a reasonably legalistic definition which stated that "a substorm is an interval of time during which energy is suddenly dissipated in association with auroral break-up." During the subsequent deliberations, the committee attempted to come to a better physical understanding of the relationship between the solar wind, the magnetosphere, and the ionosphere which is intrinsic to the overall substorm process.

It is very generally recognized that virtually all of the energy dissipated in substorms comes ultimately from the solar wind. Figure 8-1 provides a convenient summary of our present understanding of the substorm sequence. It is clear that substorm processes are a mixture of "driven" processes and of "unloading" processes. A substantial subsequent portion of this report will be devoted to clarifying the meaning of these terms. For the present, however, in Figure 8-1 it is shown that, quite generally, part of the energy input from the solar wind to the magnetosphere goes into convective dissipation in the ionosphere. This is the aspect of substorms which has been termed the "driven" component of substorms. Just as generally, however, part of the solar wind energy input goes into the form of stored energy in the Earth's magnetotail. This storage process is also "driven" by the solar wind, but the dissipation associated with the stored energy does not occur immediately.

We term the driven process "global" in Figure 8-1 since most of the magnetospheric surface, much of the magnetotail, and most of the polar cap ionospheric regions are involved in the convection and/or storage sequences. A substorm is more than just this driven process, however. To be a substorm, there has to be at least one episode of sudden (explosive) dissipation. This explosive dissipation phase represents the "unloading" of energy that was previously stored in the magnetotail. Some of the unloaded energy eventually appears in the inner magnetosphere/ionosphere in

the form of ring current enhancement, Joule heating, and auroral particle precipitation, while the rest of it is returned to the solar wind, probably in the form of a severed portion of the magnetotail (i.e., the so-called "plasmoid").

It therefore emerges quite clearly that substorms are a combination (often a very complicated combination) of driven and unloading processes. Furthermore, these processes are usually going on concurrently during the course of magnetospheric activity. It is therefore not very surprising that considerable confusion has ensued as researchers have examined different facets of substorms.

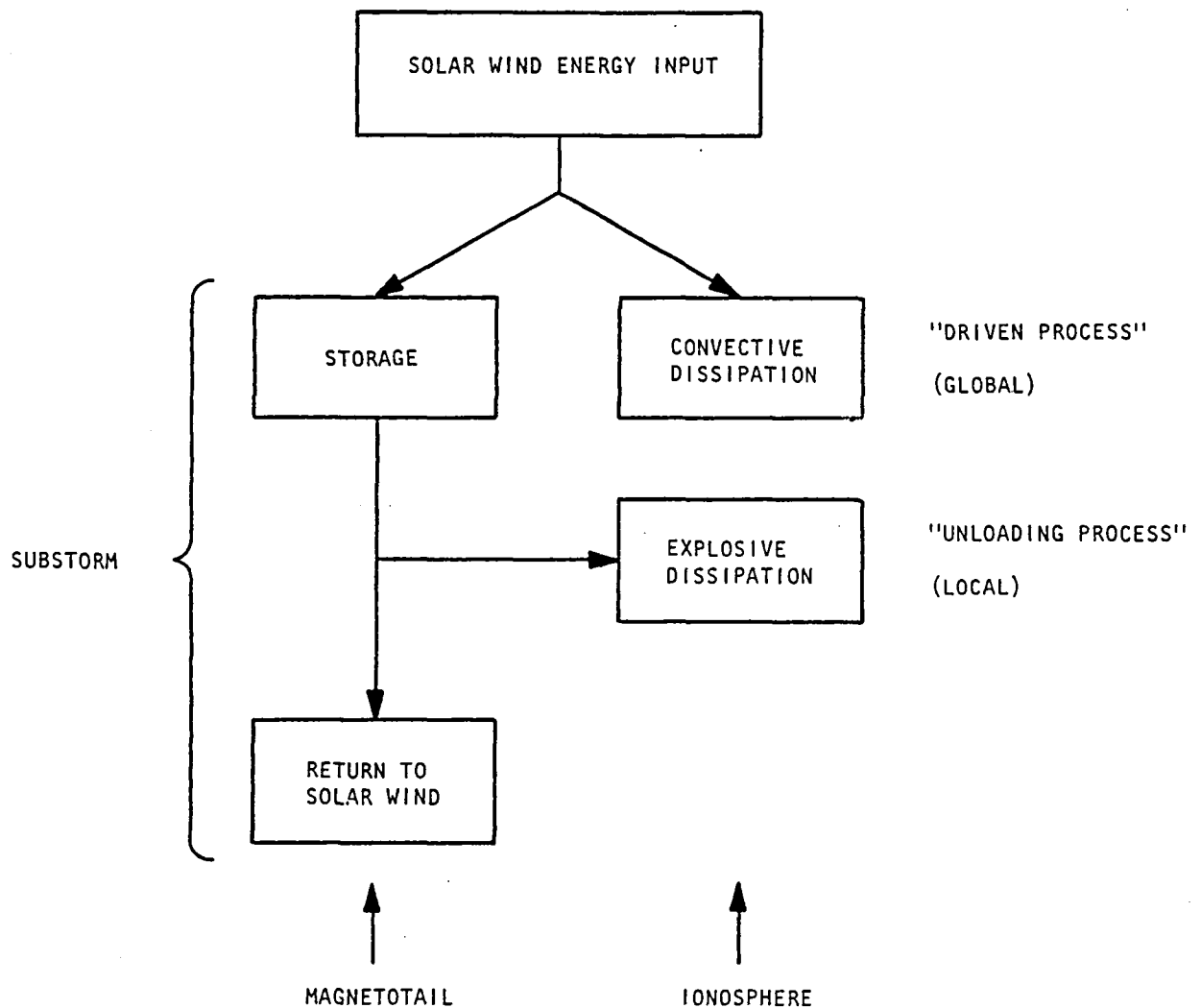


Figure 8-1. A summary of the substorm sequence showing the roles of the magnetotail and the ionosphere and also showing the aspects of substorms which are regarded as driven and unloading processes.

Based on our understanding of Figure 8-1 and based on our examination of evidence acquired throughout the solar wind-magnetosphere-ionosphere system, we arrive at the following operational definition of a substorm. A substorm comprises a set of phenomena that occur when energy entering the magnetosphere from the solar wind is dissipated in the auroral ionosphere and the ring current--partially by a process directly driven by the solar wind input and partially by one or more episodes of intermittent unloading of energy stored in the magnetotail. As an adjunct, we note that the episodes of unloading are best signified in the ionosphere by auroral expansion and are best signified in the magnetotail by plasma acceleration and jetting processes.

The above discussion illustrates that what we call the substorm process is, in reality, a large collection of physical processes. Often it is said that substorms are just a set of boxes (each box being a physical process), one placed upon another. This picture is not strictly correct. A more proper interpretation is that a substorm is a picture made up of many puzzle pieces. Each of the pieces is a physical process, but the processes must be put together in a very special way in order for the picture of magnetosphere activity (substorms) to emerge clearly.

Figure 8-2 gives a brief and very general impression of the many physical processes involved in the substorm sequence. We employ a highly simplified and schematic diagram of the solar wind-magnetosphere-ionosphere system to make our point. Solar wind energy may enter the magnetosphere by reconnection, by diffusive processes, and/or by "impulsive" entry. This energy may be directly dissipated through the action of convective electric fields and ionosphere currents, or it may be stored in the magnetotail by closely related processes. Stored energy in the magnetotail, when it is dissipated, may give rise to hot, jetting plasmas and to very energetic ions and electrons. This particle acceleration may result from reconnection processes, from parallel electric fields, or simply from the stochastic processes inherent in wave-particle interactions. Finally, the magnetosphere and ionosphere are intimately coupled (by waves, field-aligned currents, and auroral particle precipitation) and it is in the ionosphere that much of the substorm energy is eventually dissipated.

Our point in Figure 8-2 is that many detailed physical processes being discussed by other STP Workshop groups (solar wind-magnetosphere coupling, reconnection, particle acceleration, etc.) are an intimate part of substorm studies, also. These are the pieces of the substorm puzzle. It is entirely proper to study the individual parts of the overall picture in great detail. In studying substorms, however, we attempt to look at the whole picture and we attempt to understand the interrelationships, that, in large measure, constitute solar-terrestrial physics.

The study of substorms is complicated and is often fraught with controversy. Quite frequently, therefore, it is suggested that the various individual processes that putatively makeup a substorm should be studied in isolation and it is consequently argued that overall substorm studies are not very productive. Quite to the contrary, however, substorms studies are of the utmost importance in space physics. Substorms, indeed, represent a collection of many processes, but it is the assemblage of these processes in a very special way that gives rise to a substorm. Substorm studies therefore give the overview of solar-terrestrial relationships that represents the fundamental objective of our research discipline. The remainder of this report is our best current attempt to assess where we stand in this enterprise and, also, is our estimation of where we should move next in our continuing effort to understand the dynamics of near-Earth space.

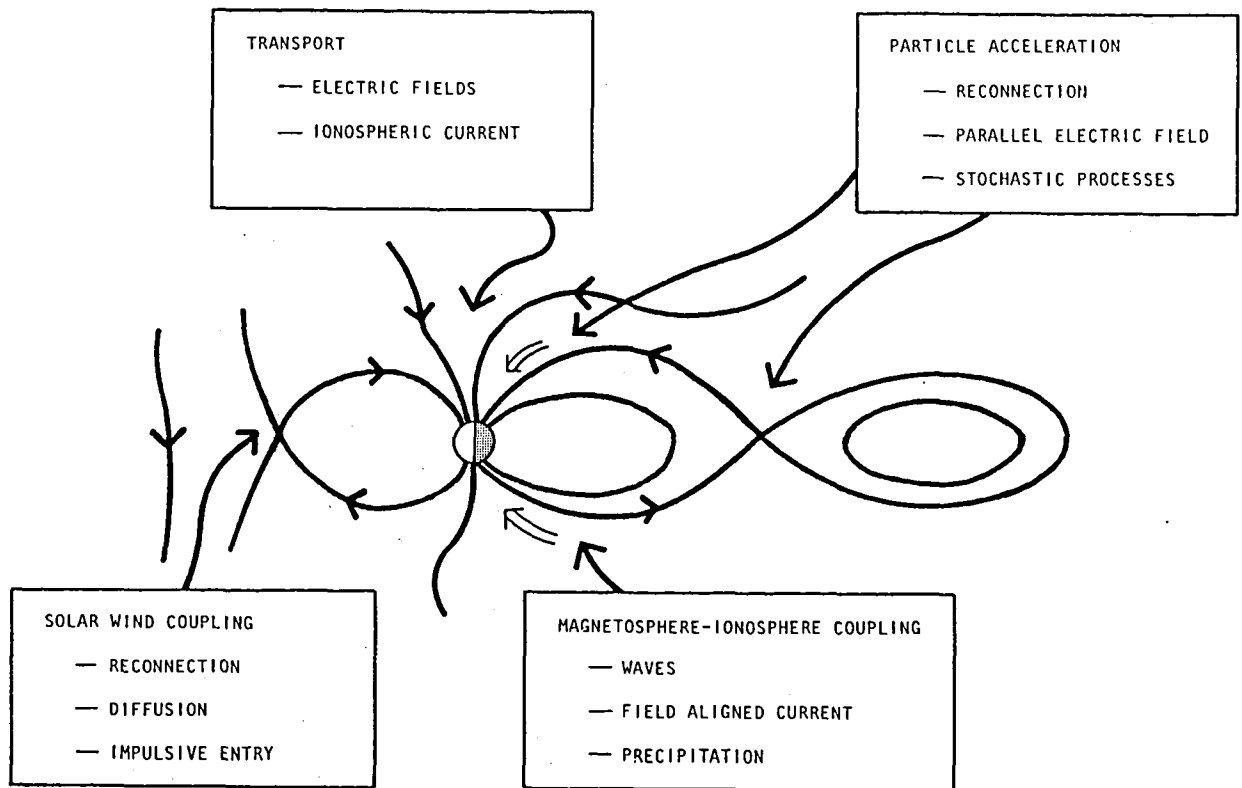


Figure 8-2. A highly schematic cross-sectional view of the magnetosphere which illustrates some of the physical processes which are believed to be involved in substorms.

## II. SOLAR WIND - MAGNETOSPHERE COUPLING

The terrestrial magnetosphere is highly dynamic with large-scale variations occurring on the time scales of minutes to days. It has long been recognized that the energy to drive these magnetospheric processes comes from the solar wind. The energy flux ( $\text{ergs cm}^{-2} \text{s}^{-1}$ ) in the solar wind that supplies the energy requirements of the magnetosphere could be derived (in a very elementary and dimensionally consistent sense) by one of two means. The first is the solar wind particle kinetic energy flux which corresponds to  $\sim 10^{-1} - 10^2 \text{ erg/cm}^2\text{-s}$  at the Earth's orbit. The other source would be the magnetic energy flux which corresponds to  $\sim 10^{-4} - 10^{-1} \text{ erg/cm}^2\text{-s}$  for the typical range of solar wind parameters that occur near the Earth.

Given these considerations, we note that substorm studies have progressed from a descriptive study of global auroral morphology in the early 1960s to the present quantitative study of the solar wind-magnetosphere-ionosphere system. In recent years, correlation studies of solar wind quantities with magnetospheric quantities have much improved our understanding of magnetospheric substorms. Most recently, solar wind quantities, or their combinations, such as  $V$ ,  $\rho V^3$ ,  $B_z V$ ,  $V^2 B_z$ ,  $\epsilon$ , etc., have been extensively correlated with various geomagnetic indices, or their combinations, and theoretical implications of the results have been examined in terms of the solar wind-magnetosphere coupling and the magnetosphere-ionosphere coupling (see Table 8-1).

The details of physical mechanisms for coupling the solar wind to the magnetosphere are the purview of Working Group 5. We refer the reader to that group's report for explicit discussion of various coupling processes. In this report we concentrate on the conditions in the solar wind which are known to give rise to enhanced geomagnetic (substorm) activity.

### INTERPLANETARY PARAMETER CHANGES

Following the theoretical suggestion of Dungey (1961), the experimental evidence accrued over the past two decades strongly supports the interpretation that solar wind-magnetospheric coupling is of an electromagnetic character. As first shown by Fairfield and Cahill (1966), the north-south component ( $B_z$ ) of the interplanetary magnetic field (IMF) is of fundamental importance: a southward IMF ( $B_z < 0$ ) gives rise to enhanced magnetospheric activity, while a prolonged northward IMF ( $B_z > 0$ ) is followed by quiet geomagnetic conditions.

Numerous subsequent studies have further established and clarified the important coupling role played by the IMF north-south component (e.g., Arnoldy, 1971; Foster et al., 1971; Tsurutani and Meng, 1972; Meng et al., 1973). Although the solar wind speed,  $V$ , tends to be substantially less variable than  $B_z$ , many studies have also shown that  $V$  is correlated with magnetospheric activity (e.g., Snyder et al., 1963). Thus, the "combined" interplanetary parameter  $VB_z$  has been widely used as an index of coupling between the solar wind and the magnetosphere (Aubry and McPherron, 1971; Rostoker et al., 1972; Holzer and Slavin, 1978; Meloni et al., 1982).

### "BEST" SUBSTORM PREDICTORS IN THE SOLAR WIND

It has become a matter of increasing concern and debate whether a "best" interplanetary coupling parameter can be specified and, whether improved physical understanding of the coupling mechanism can be obtained by comparing various solar wind parameters to geomagnetic activity changes. Burton et al. (1975), for example, presented evidence that the magnetosphere acts as a half-wave rectifier responding readily to  $B_z < 0$ , but exhibiting little response to  $B_z > 0$ . It was therefore suggested that the interplanetary quantity  $VB_s$  (where  $B_s$  is the southward IMF component) should be used as the best coupling parameter and, in relation to the physics of the

TABLE 8-1  
LIST OF THE CORRELATION STUDIES BETWEEN GEOMAGNETIC INDICES  
AND SOLAR WIND PARAMETERS

Author(s)	Geomagnetic index (Time resolution)		Relation with solar wind parameters
Synder et al. (1963)	$\Sigma Kp$	(24 hr)	$\Sigma Kp = (V - 330)/8.44$
Olbert (1968)	$\Sigma Kp$	(24 hr)	$\Sigma Kp = (V - 262)/6.3$
Ballif et al. (1969)	$Kp$	(3 hr)	$Kp = 9 \left[ 1 - \exp \left( -\frac{\sigma B_{TN} - 0.35}{7.70} \right) \right]$
Arnoldy (1971)	AE	(1 hr)	$AE = -0.26(\Sigma B_s \tau)_0 - 0.91(\Sigma B_s \tau),$ $= -0.33(\Sigma B_s \tau)_2 + 0.12 P_0$
Bobrov (1973)	$Kp$	(3 hr)	$Kp = f(V, B_z, \Delta B_z)$
Garrett et al. (1974)	ap, AE	(1 hr)	$ap, AE = C_1 + C_2 V(B_s) + C_3 V\sigma$
Murayama and Hakamada (1975)	AE	(1 hr)	$AE = C B_s V^2$
Burton et al. (1975)	Dst	(2.5 min)	$\frac{d}{dt} Dst_0 = F(E) - a Dst_0$
Crooker et al. (1977)	Ap		$Ap = 3.5 \times 10^{-5} B_z V^{2-1.9}$
Svalgaard (1977)	am	(3 hr)	$am = 6.6 q(f, \alpha) \frac{BV_0}{21} \left[ \frac{nV_0^2}{105} \right]^{1/3}$ $\times \frac{1.157}{(1+3 \cos^2 \psi)^{2/3}}$
Maezawa (1979)	AL, AU, am	(1 hr)	$AL \propto B^{0.85} V^{2.08} (\sin \theta)^{0.54}$ $AU \propto B^{0.67} V^{1.15} (\sin \theta)^{0.34}$ $am \propto B^{1.03} V^{2.34} (\sin \theta)^{0.37} n^{0.2}$
Holzer and Slavin (1979)	AL	(1 hr)	Rate of erosion $\frac{d\phi_e}{dt} = 0.2(1.9 \times 10^{10} \text{ cm}) B_z V_{sw} (\text{km s}^{-1})$ Rate of return $\frac{d\phi_r}{dt} = 1.8 \times 10^{10} AL(\gamma)$
Murayama (1979)	AL	(1 hr)	$AL = 60(B_s + 0.5) V^2 F(\chi, B_y) n^{0.13}$
Murayama et al. (1980)	AL	(1 hr)	$AL/V^2 = K(B_s + 0.5)$
Baker et al. (1981)	AE	(2.5 min)	$VB_z, V^2 B, B^2 V, \text{etc.}$
Lei et al. (1981)	Correlation between the interplanetary and magnetospheric electric fields		
Reiff et al. (1981)	Correlation between $\phi_{PC}$ and $B_z, \epsilon$		
Holzer and Slavin (1982)	AL		$BV^2, B_s V, B_s^2 V$
Iijima and Potemra (1982)	Correlation between field-aligned current density and $VB_z, V^2 B, B^2 V, \text{etc.}$		
Meloni et al. (1982)	AE		$B_z, VB_z, \epsilon$

Note: For details of the expressions and definitions of each notation, see the references in Akasofu (1981).



coupling, these results suggested the importance of the dawn-to-dusk electrostatic field as a means of transferring energy to the magnetospheric system.

Subsequent to the work of Burton et al., several other studies showed the importance of the interplanetary ( $-V \times B$ ) electric field in the development of storms (e.g., Burlaga and Lepping, 1977) and substorms (e.g., Caan et al., 1978). It was further found that general magnetospheric activity (as measured by Kp or Ap) and substorm activity (as measured by AE and/or AL) was even better correlated with  $V^2 B_s$  than with  $VB_s$  (Crooker et al., 1977; Murayama and Hakamada, 1975; Murayama et al., 1980; Holzer and Slavin, 1982).

A different formulation of the solar wind-magnetosphere coupling parameter was presented by Perreault and Akasofu (1978). This formulation suggested that

$$\epsilon = VB^2 \sin^4(\theta/2) \ell_0^2$$

is the best predictor of total magnetospheric energy output; here  $\theta$  is given by  $\tan^{-1}(B_y/B_z)$ , or by this angle's supplement for  $B_z < 0$ ,  $\ell_0$  is a constant ( $=7 R_E$ ). Akasofu (1979) suggested that  $\epsilon$  is also the best interplanetary predictor of individual substorms as measured by AE.

Note that although  $\epsilon$  is roughly proportional to  $VB_s^2$ , it has a finite value for  $B_z > 0$ , whereas  $B_s$  is defined as

$$B_s = \begin{cases} 0 & \text{for } B_z \geq 0 \\ -B_z & \text{for } B_z < 0 \end{cases}$$

Perreault and Akasofu (1978) and Akasofu (1979) strongly emphasized the difference between  $\epsilon$  and parameters involving  $B_s$  and discussed the  $\epsilon$  function in the prediction of (and correlation with) magnetospheric storm and substorm parameters.

Many of the studies referred to in the foregoing discussion have involved low-time resolution data, such as one- and three-hour averages of solar wind and magnetospheric parameters (e.g., Arnoldy, 1971; Holzer and Slavin, 1978, 1982; Akasofu, 1979; Murayama et al., 1980). Since magnetospheric substorms tend to develop on time-scales of one hour or less, significant time aliasing and masking of physical effects is possible with long-term average data. Recent high-resolution studies (Baker et al., 1981a, 1983; Clauer et al., 1981) have compared the various interplanetary parameters and have shown general similarities in the capabilities of  $\epsilon$  and  $VB_s$  for prediction of substorms (as measured by AE).

As implied by many of the studies referred to in the foregoing, much of our understanding of "best" coupling parameters relies upon the statistical use of various geomagnetic indices such as AE (the auroral electrojet index). However, the accuracy and limitation of all indices should always be borne in mind. The AE index, for example, is not strictly speaking a substorm index. As pointed out by Kamide and Akasofu (1983), the AE index does not properly measure the total electrojet current, but rather it is an index of current density. Likewise, a number of assumptions used in associating the Dst index with the total ring current kinetic energy should be carefully examined.

We fully expect that new advances in our understanding of energy coupling within the solar wind-magnetosphere-ionosphere system will come about through significant improvement in our quantitative estimates of such global quantities as the ring current injection rate, the Joule heat production rate, tail energy dissipation rate, etc. Since these estimates presently rely on geomagnetic indices (for want of other generally usable methods), we must continually strive to improve the continuously available indices. The indices used, now and in the future, are measures of the underlying geomagnetic disturbance and, therefore, we must improve and extend our understanding of the relationship of indices to the physical disturbance fields.

### TIME DELAYS: SOLAR WIND - MAGNETOSPHERE COUPLING RESPONSE TIME

The response of the magnetosphere to the solar wind has been studied by several techniques which provide information about the time constants of the system. These techniques include superposed epoch analysis, cross-correlation analysis, measurement of time delays, and linear prediction filtering. All of these techniques indicate that the magnetosphere responds to the solar wind with delays of order 20-60 minutes. For example, one of the first studies of response time was the work of Foster et al. (1971), who used superposed epoch analysis to show that isolated substorm expansion onsets are preceded by an hour of southward interplanetary magnetic field. Caan et al. (1978) in a similar study demonstrated that the tail magnetic field and mid-latitude magnetic disturbance near midnight are also correlated with expansion onsets. Their results showed that major expansion onsets are preceded by an increase in tail magnetic fields of approximately one hour duration.

Another study was that of Meng et al. (1973) who determined the correlation function for AE versus interplanetary  $B_z$ . Their results are similar to those of a more recent general study of Baker et al. (1981a), who found that AE peaks approximately 40-60 minutes later than parameters such as  $VB_s$  (rectified electric field),  $V^2B_s$ , and  $\epsilon$ .

In yet another technique which uses case histories, various investigators have found delays of 30-90 minutes between southward turnings of the IMF and sudden changes in phenomena near midnight. For example, McPherron et al. (1973) obtained 32 minutes for this delay. Baker et al. (1981b), in a comprehensive study of a large isolated substorm found ~90 minutes between the beginning of a tail field increase and a major expansion onset. Iyemori et al. (1979) found an average delay of ~60 minutes between southward turnings and the first observable Pi2 of a substorm.

The technique of linear prediction filtering introduced by Iyemori et al. (1979), provides the most detailed information about magnetospheric response time presently available. In this technique, time series of observed solar wind input and magnetospheric output are used to construct the most general possible relation between the two time series. This relation is represented as a time domain filter which may be convolved with an arbitrary observed input time series to predict the system output. Using high time resolution data Clauer et al. (1981, 1983) determined filters relating the dawn-to-dusk solar wind electric field  $VB_s$  to the AL and ASYM indices. The two filters are very similar and have the form of a Rayleigh pulse peaking at 40-60 minutes and decaying to background after two hours. More recently, Bargatze et al. (1983), have calculated filters as a function of the level of activity. Their results suggest the magnetospheric response is made up of two parts. The first part peaks near 60 minutes and is dominant during low levels of activity. The second part peaks near 20 minutes and is dominant during disturbed times. The authors suggest that the short response time represents the driven component of AL activity while the larger time represents unloading of energy from the tail.

The success of the linear prediction filters in predicting a substantial fraction of the variance in the AL and ASYM indices (~40%) appears to suggest that geomagnetic activity is driven by the solar wind, albeit with some delay and with frequency-dependent attenuation of the input. It should be noted, however, that these results cannot answer detailed questions of whether a specific energy release mechanism plays a significant role in the dissipation of energy extracted from the solar wind. For instance, as discussed below, energy storage in the tail is driven by the solar wind through flux transport into the tail. If in addition, the release of energy is also driven, i.e., deterministically related to the input, then geomagnetic activity as measured by the AL index will be predictable using solar wind data alone. Other information, for example, the observed sequence of phenomena in the geomagnetic tail with sudden increases in AL must be cited to demonstrate the role of energy storage/release in substorms.

### III. DRIVEN PROCESSES IN SUBSTORMS

As discussed in the Introduction, it is now recognized that substorms are made up of a complex combination of driven and unloading processes. A driven process is by definition a global phenomenon encompassing much of the magnetosphere which is deterministically related to the solar wind. To the extent that the magnetosphere may be approximated by a linear system, driven processes may be represented by an impulse response function. If this function is known, convolution of it with a measured input enables one to predict the output. For an impulse response that approximates a delta function at some delay the output will be nearly identical to the input but time shifted. As shown by Akasofu (1981a), the time series of hourly averages of the total energy parameter  $U_T$  and AE index are very similar to the solar wind energy input parameter  $\epsilon$ . Figure 8-3 illustrates this point with a four day interval including a moderate magnetic storm. High correlations between input and output time series at this resolution have led Akasofu (1981a) to suggest that the magnetosphere is primarily a directly driven system.

There are two major driven processes involved in substorms. The first is dissipation. Dissipation occurs in a variety of ways. The most important forms of dissipation include Joule heating by ionospheric currents, particle precipitation into the polar regions, and particle injection into the ring current. Ring current injection is actually a storage process except that energy stored in this way is not available to drive substorms. Other means for dissipation which may be important include ultra low frequency waves and auroral kilometric radiation.

The second driven process in substorms is storage. Energy is stored in magnetic fields surrounding the currents connected to the ionosphere and in the currents of the magnetotail. Energy is also stored in the electric field produced by enhanced drift of convecting particles.

#### DISSIPATION

As discussed in Section II, the three major dissipation processes depend on the magnitude of the solar wind velocity and magnetic field, and on the angle the field makes relative to the Earth's dipole axis. In general, the strength of the dissipative processes depends on the velocity and field magnitudes, but their temporal behavior depends on the field orientation. The strong dependence of the driven processes on solar wind field orientation has suggested to most workers that magnetic merging on the sub-solar magnetopause is their primary cause. An important consequence of magnetic merging is that a fraction of the solar wind electric potential is applied across the magnetosphere. In the simplest approximation this fraction corresponds to the sheath electric field times the length of the merging region on the dayside magnetopause.

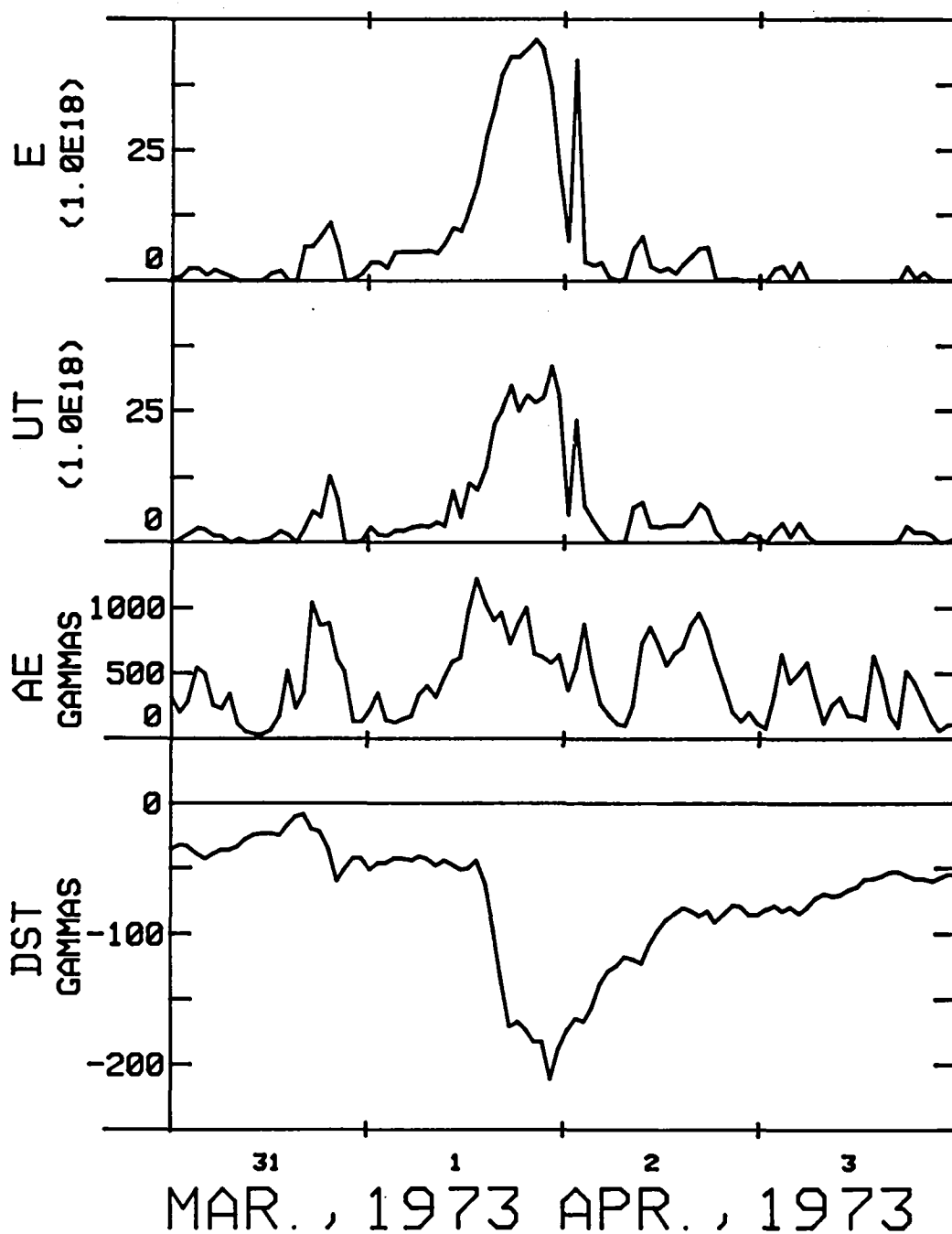


Figure 8-3. An example of the relationship between the power  $\epsilon$  of the solar wind-magnetosphere dynamo and the total energy dissipation rate  $U_T$  of the magnetosphere for the storm of March 31, - April 3, 1973. Both the AE and Dst indices are also shown.

At the edges of this region magnetosheath field lines are swept past the magnetosphere. Within the interaction region they merge with the magnetosheath field and are transported across the polar cap and into the tail lobes. Since the field lines are good conductors the potential across the merging region is transferred to the polar cap. Eventually these same field lines define the boundary of the magnetotail, and the merging potential appears across the tail as well. Because the width of the tail is greater than the length of the merging region the actual potential across the tail is a fraction of the solar wind electric field present over the same distance in the undisturbed solar wind.

The above process has been called the solar wind dynamo (Akasofu, 1981a) because a conductor, viz., the magnetosheath plasma, moves across a magnetic field, i.e., the geomagnetic field lines connected to the solar wind. Lorentz forces acting on the charges in the conductor separate charges which accumulate on the magnetopause, positive on the dawn side and negative on the dusk side. The potential difference across the tail drives the field-aligned currents which dissipate energy in the process of Joule heating. A circuit analogy of this driven process is sketched in Figure 8-4. Recent work by Reiff et al. (1981) reports a quantitative relation between the polar cap potential and  $\sqrt{\epsilon}$ , i.e., to the magnitude of the solar wind magnetic field (or electric field) consistent with this simple explanation.

### Solar Wind-Magnetosphere Dynamo

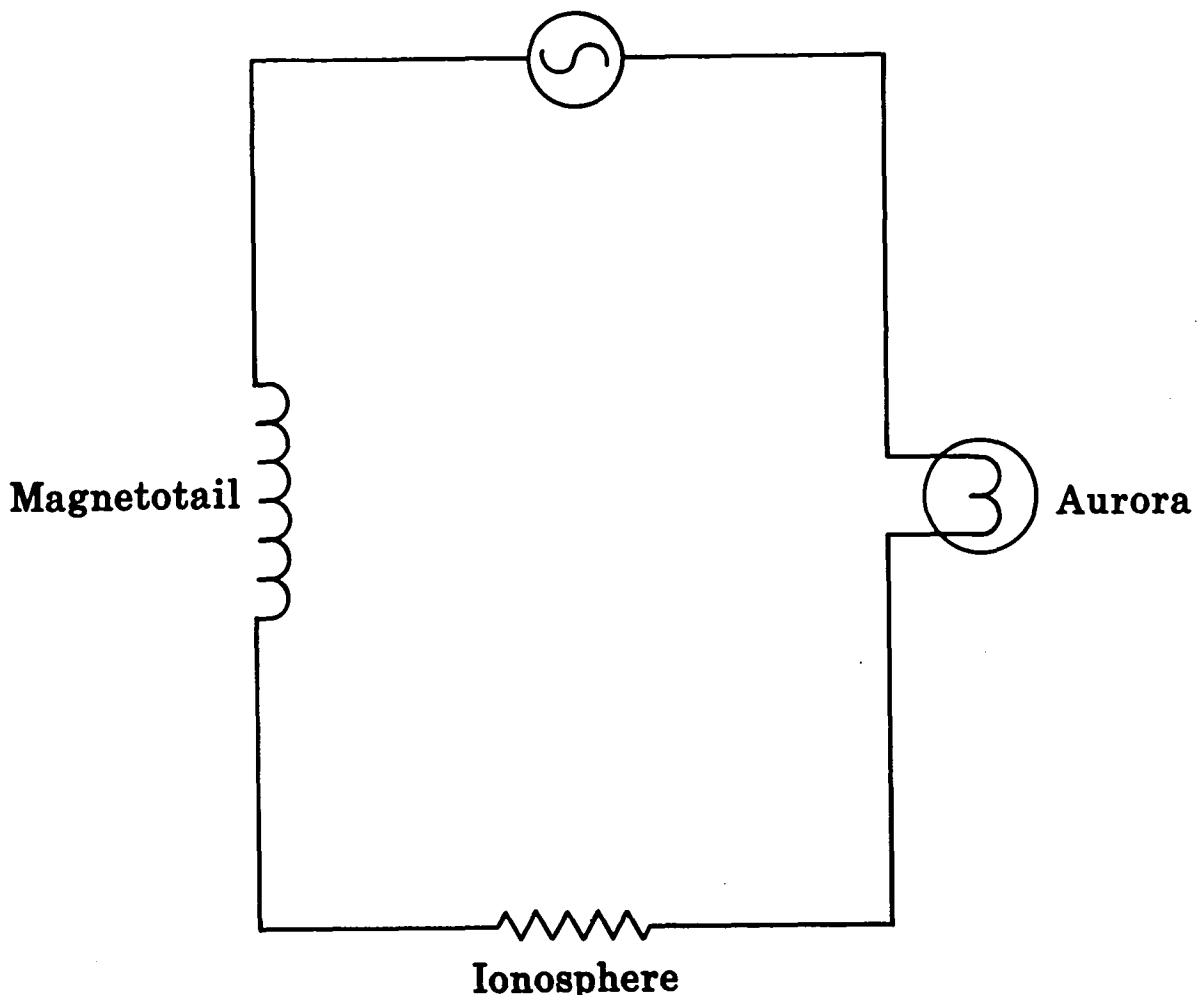


Figure 8-4. The equivalent circuit of a directly driven magnetospheric model in its simplest representation.

### Joule Heating

As mentioned above, magnetic field lines emerging from the poleward edge of the auroral oval are connected to the space charges created by the interaction of the solar wind with the magnetosphere. Since these field lines are extremely good conductors, they transmit the solar wind potential to the ionospheric boundaries of the polar cap. This is illustrated schematically in Figure 8-5 by representative field lines in the dawn-dusk meridian plane. The ionosphere, however, is a resistive conductor and closure of the field-aligned currents results in the dissipation of solar wind energy through Joule heating.

Closure of the field-aligned currents is more complex than one might first think. The polar cap, particularly in winter, is a very bad conductor as compared to the auroral oval. Furthermore, the magnetospheric electric field when projected onto the polar auroral oval lies more in meridian planes than azimuthal circles. The consequence is that most of the current does not close across the polar cap.

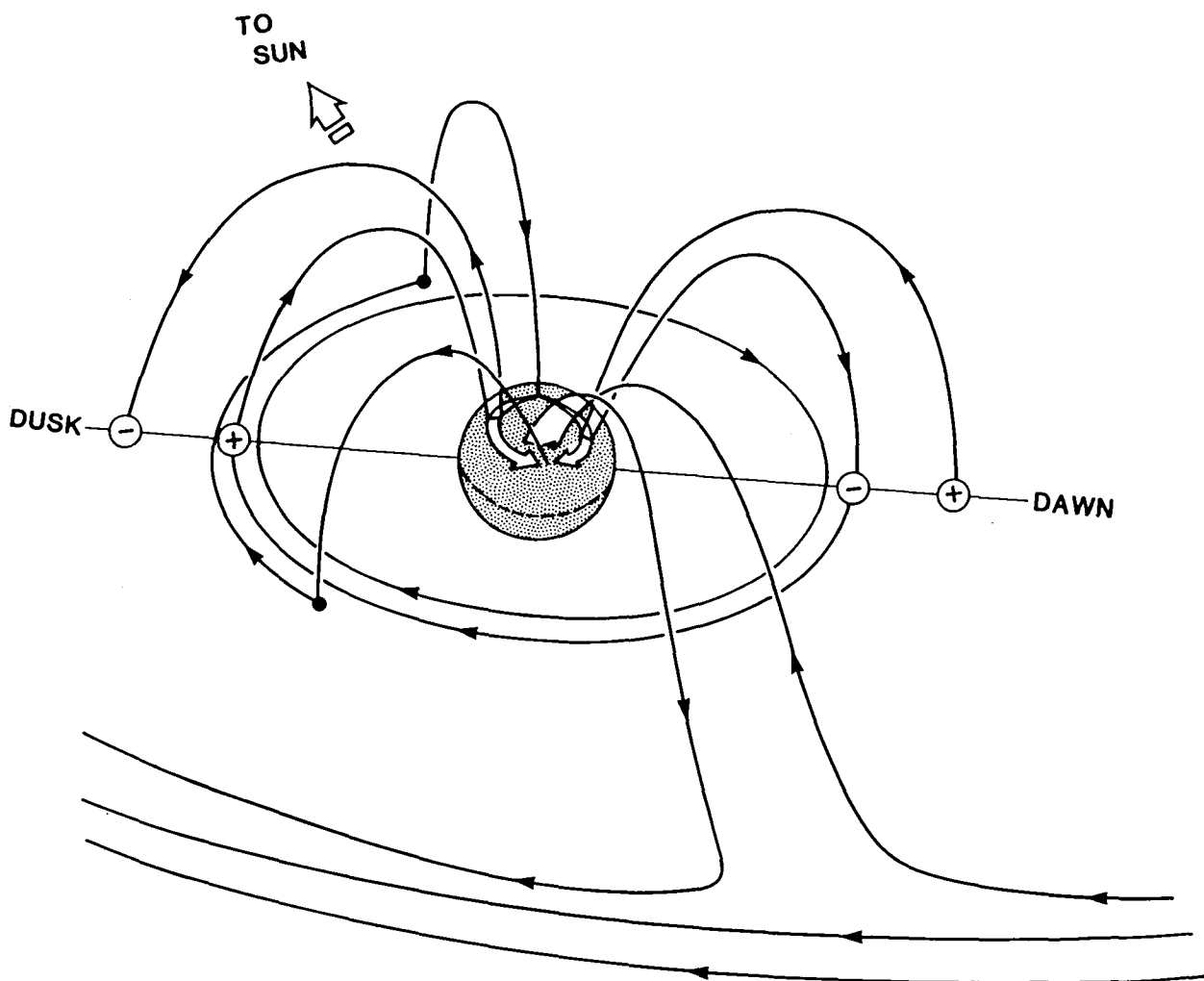


Figure 8-5. Schematic drawing showing representative field lines in the dawn-dusk meridian plane.

A partial explanation for the observed electric field orientation is the development of an internal space charge within the magnetosphere which "shields" the inner region close to the Earth from the effects of the solar wind electric field. This shielding charge is created by the tendency for particles drifting sunward under the influence of the external electric field to separate due to curvature and gradient drift in the Earth's magnetic field. Electrons from the tail drift inward and towards the dawn side, while protons drift toward dusk. The space charge created by these drifts is opposite to that created by the solar wind electric field. As evident from the diagram, the ionospheric electric field on the dawn side is equatorward, while that on the dusk side is poleward.

The field-aligned currents driven by the solar wind in the dawn sector thus flow down field lines on the poleward side of the auroral oval and south across the oval. They then flow upward at its equatorward edge. On the dusk side the currents are in the opposite direction. The circuit is closed by the opposing drifts of electrons and protons in the night sector which create a partial, equatorial ring current from dawn to dusk across the nightside of the magnetosphere.

### Particle Precipitation

The second major dissipative process driven by the solar wind is particle precipitation. On the night side, particles drifting earthward from the tail under the influence of the solar wind electric field are scattered into the loss cone by waves generated near the equatorial plane. The waves are produced by plasma instabilities driven by anisotropies and gradients in the particles. These, in turn, are created by the changing magnetic field encountered by the particles as they drift earthward. The poleward edges of the regions of precipitation are thought to be defined by the last closed field which define the upper and lower boundaries of the plasma sheet. The equatorward edges, on the other hand, are defined by the location of a balance between earthward convection and wave scattering into the loss cone. The ionospheric region between the two boundaries constitute the nightside auroral oval.

On the dayside, particles from the magnetosheath are scattered onto field lines emerging from the polar cusp. Particles with initially small pitch angles and those scattered to small angles by instabilities are lost to the ionosphere. The region where these particles enter the dayside ionosphere maps into the dayside portion of the auroral oval.

On the nightside a third process is very important in particle precipitation. This is the formation of field-aligned potential drops. These potential structures are formed where upward field-aligned currents exceed a critical density of order  $10^{-6} \text{ A/m}^2$  (Akasofu, 1981c). The effect of the potential drop is to accelerate electrons toward the ionosphere and ionospheric ions away. Electrons precipitated into the ionosphere by this process have sufficient energy to create discrete arcs of considerable luminosity. Since these structures can form suddenly they may account for sudden brightening of the aurora prior to the release of energy stored in the tail magnetic field.

It has also been suggested, (Akasofu, 1981c) that this process can lead to positive feedback. Enhanced ionospheric conductivity produced by the accelerated electrons, leads to increased ionospheric currents, which in turn require enhancement of the field-aligned currents which feed them. This then requires a larger potential drop which accelerates electrons even more. As discussed below in Section IV, it is this instability of the driven current system which is used to explain the onset of the substorm expansion phase (Akasofu, 1981c).

Precipitation of particles into the ionosphere has two major consequences. First, through ionization, it creates the conductivity which channels the ionospheric currents and is responsible for Joule heating. Second, the energy carried by the precipitating particles is dissipated in the atmosphere. Since these processes of precipitation are primarily controlled by the large scale magnetospheric electric field, they may be classified as driven processes and are expected to depend on the strength of the solar wind power input.

### Ring Current Injection

The third process which dissipates solar wind energy is ring current injection. As mentioned above, from the point of view of the solar wind, the ring current is a temporary reservoir for storage of energy. This energy is eventually dissipated in the atmosphere, primarily by particle precipitation. The time scale for this dissipation is quite long (approximately 8 hours) compared to a typical substorm. Furthermore, this energy is not available to drive substorms. Consequently, from the point of view of substorms, solar wind energy is dissipated if it is injected into the ring current.

The idea that ring current injection is driven by the solar wind is not universally accepted. Many investigators feel that ring current development is primarily a consequence of impulsive particle injection during substorm expansion phases. However, the high correlation between solar wind parameters such as  $VB_s$  or  $\epsilon$  and Dst suggests there is a large driven component to ring current injection. Resolution of this point is an important outstanding problem of magnetospheric physics.

A driven model for ring current injection follows the ideas of Wolf and coworkers (Harel et al., 1981). Under quiet conditions the magnetosphere approaches a steady state of low-level convection. Particles drift earthward under the influence of low-level electric fields. Gradient and curvature drifts take over producing a westward ring of current across the nightside. Precipitation processes produce a relatively sharp inner edge to the drifting particles. At quiet times, most of the particles drift into the magnetosheath before crossing the noon meridian.

When the solar wind electric field across the magnetosphere is enhanced by a southward IMF the situation changes. The precipitation boundary moves earthward. Particles approach closer to the Earth and gain greater amounts of energy drifting across the enhanced magnetospheric electric field. Furthermore, many particles cross the noon meridian. More particles of higher energy drifting closer to the Earth, and nearly circling it, create an enhanced ring current.

If the magnetospheric electric field were to remain constant, particles would lose energy as they drift across the dayside. However, usually the IMF turns northward suddenly. Particles which would have eventually lost energy and returned to the magnetosheath are suddenly trapped on closed drift shells. These particles rapidly distribute themselves around the Earth in a symmetric ring of current. Eventually, wave particle interactions scatter particles into the loss cone. Charge exchange between these scattered particles and atmospheric particles leads to a decay of the ring current over a period of several days.

As will be discussed below there also appears to be an impulsive component of ring current injection. This component is related to the unloading process in which energy stored in the tail is released explosively during the expansion phase of a substorm. Particles energized by a variety of processes involving induced electric fields are injected in a localized region near midnight. Their subsequent drift augments the ring current created by the driven process.



## STORAGE

The driven process is also responsible for storage of energy in the magnetotail. In contrast to the ring current situation, this energy is available for substorms. As discussed in a subsequent section, the release of this energy by the unloading process contributes significantly to substorm disturbances.

Energy is, of course, also stored in magnetic fields around the currents of the driven system. However, to the extent that the geometry of the currents and the conductivity of the auroral oval remain unchanged, the time constant for release of this energy will be the same as the time constant for its storage. This assumption is essential to the concept of a driven system which presupposes the input to the system is transformed by passive elements into output (dissipation).

## Tail Field Enhancement

As mentioned above, the correlation of geomagnetic activity with a southward IMF strongly suggests that the solar wind interaction with the magnetosphere involves merging of geomagnetic and interplanetary magnetic fields. An important consequence of merging is that the solar wind transports merged field lines across the polar cap into the lobes of the magnetotail. These field lines are open to the solar wind and are the means whereby a fraction of the solar wind electric field is applied across the magnetosphere. These lines must eventually reconnect and return to the dayside. Otherwise, the entire dayside magnetosphere would be eroded away to become polar cap field lines. This reconnection is thought to occur initially at a distant neutral line far down the tail. Once reconnected, the field lines drift earthward eventually returning to the dayside.

If the rate of reconnection at the distant neutral line were to adjust itself immediately to match the dayside, the flux in the tail lobes would remain constant. Experimentally however, this is not observed. Instead, the polar cap expands as magnetic flux accumulates in the tail. This implies a gradual enhancement of the tail current. Also, as mentioned above, the precipitation/flow boundary at the inner edge of the plasma sheet moves earthward in response to the enhanced convection electric field. These two effects cause an increase in the inclination of the field at synchronous orbit. This significantly effects the pitch angle distribution of energetic electrons drifting through the night sector, forcing them to become more field aligned. This in turn increases the probability they will be lost to the atmosphere through precipitation.

## Plasma Sheet Changes

An imbalance between dayside and nightside reconnection also effects the plasma sheet. Closed field lines within the plasma sheet convect earthward to provide the flux needed in the dayside merging region. Since less flux is being reconnected on the nightside than required on the dayside the plasma sheet must thin. Observationally it appears that the most rapid thinning is close to the Earth. This thinning in conjunction with the growth of a tail-like field appear to be the condition necessary for release of energy stored in the magnetic field of the tail lobes. As discussed in a later section there seem to be a variety of mechanisms which trigger the release of this energy.

Unbalanced flux transfer also significantly affects the dayside magnetopause. Since flux does not flow into the dayside region as fast as it is transported to the tail the magnetopause is eroded. This causes the location of the merging region to move earthward and the polar cusps to move equatorward. The dayside auroral oval which is the projection of the polar cusp thus also moves equatorward. In addition, particles of 90° pitch angle which would normally drift to synchronous orbit at midnight are lost to the magnetosheath. Their absence at midnight contributes to the growth of "cigar shaped" particle angle distributions.

#### IV. THE UNLOADING PROCESS IN SUBSTORMS

As pointed out above, solar wind energy input plays a crucial role in both the driven and unloading models of substorms. As noted in the Introduction, however, it is central to our working definition of substorms to have a quality of suddenness which, in turn, corresponds to the expansion phase of substorms. It is this sudden, nearly explosive onset of energy dissipation which has traditionally been taken (Akasofu, 1964) as marking the onset of substorm activity. In this section we examine results which have established this sudden energy release component of substorm processes.

##### ONSET

The view of substorms that emphasizes the role of the magnetotail cites evidence that the magnetotail approaches an unstable state prior to onset. The continued storage of energy in the magnetotail is frequently interrupted by the sudden release of this energy at the onset of the expansion phase of the substorm. When, where, and why this onset occurs are fundamental problems of substorm research.

The preceding description of the processes beginning immediately after a southward IMF turning apply equally well to the driven and unloading models of substorms. In both models flux accumulates in the magnetotail. In the driven model it is assumed that the only consequence of this accumulation is an increase in the size of the polar cap. The unloading model assumes there are other consequences. One of these is the development of a tail-like field in the night sector. This development in conjunction with the enhanced electric field, alters the trapped magnetospheric particle drift paths. These changes are seen in the auroral ionosphere as changes in the location and character of particle precipitation. In particular there is an increase in energy dissipation due to particle precipitation which accompanies the increase in Joule heating mentioned above.

Another consequence of the imbalance of flux transfer is a thinning of the near-Earth plasma sheet. Enhanced convection within the inner magnetosphere moves plasma initially in the night-side plasma sheet toward the dayside, thinning the plasma sheet. Alternatively, instead of flowing into the inner magnetosphere, the plasma may be stored intermediately in the magnetotail by time-dependent convection (Schindler and Birn, 1982; Birn and Schindler, 1983). In the growth phase description of substorms it is argued that these changes create a progressively more unstable magnetotail. Eventually, a perturbation such as a solar wind dynamic pressure change, a northward turning of the IMF, or other, as yet unidentified process, triggers the release of the energy stored in the tail. As discussed in a subsequent section this release is manifested as a sudden diversion of a portion of the tail current through the ionosphere as a westward electrojet.

##### Internal Trigger

Data from spacecraft in geostationary orbit show a remarkably consistent sequence of variations of the electron anisotropy before and during magnetospheric substorms. For periods typically 1-2 hours prior to the onset of substorms, distributions of 10-100 keV electrons in the pre-midnight sector are observed to be peaked along the direction of the local magnetic field. These cigar-like anisotropies are accompanied by a local tail-like magnetic field which may develop further during the event. At substorm onset an abrupt transition usually occurs from the cigar-shaped distributions to pancake-shaped distributions (Baker et al., 1978). This anisotropy sequence may be due to the buildup and subsequent release of stresses in the magnetotail; the cigar phase may also be due to associated processes (e.g., compression) at the dayside magnetopause causing a loss of 90° pitch angle particles. Irrespective of the mechanism, these electrons

appear to show the cigar-shaped distributions at  $6.6 R_E$  if, and only if, the magnetosphere is headed toward the major energy release of a substorm. If no substorm is impending, the low-energy electrons simply drift through the midnight sector maintaining near-isotropy. Such data suggest that these electron anisotropies are intimately related to the growth phase of magnetospheric substorms (Baker et al., 1978; 1979a, b) and are suggestive of the buildup of internal magnetospheric stresses prior to expansion onset.

An instability in the stressed magnetotail seems to be a likely means for releasing energy at substorm onset. Substantial theoretical work in the last two decades has focused on the collisionless tearing mode as the plasma instability that would give rise to a macroscopic neutral line (cf., Section IV) in the nightside plasma sheet (Coppi et al., 1966; Schindler, 1974; Galeev and Zeleni, 1976; Galeev et al., 1978; Terasawa, 1981; Goldstein, 1981). For realistic electron and ion thermal velocities and for realistic magnetic field configurations in the Earth's plasma sheet, the relative behavior of electrons and ions would be as illustrated in Figure 8-6. Because of their very small gyroradii, thermal electrons remain adiabatic and in their bounce motion move gyroscopically through the quasi-neutral sheet (Figure 8-6a). On the other hand, thermal ions have much larger gyroradii, and for sufficiently small field line radius of curvature and sufficiently large gyroradii (Figure 8-6c) ions will become unmagnetized near the neutral sheet. The nonadiabatic behavior of the ions can lead to the ion tearing mode instability with the resultant growth of neutral line structures (Schindler, 1974).

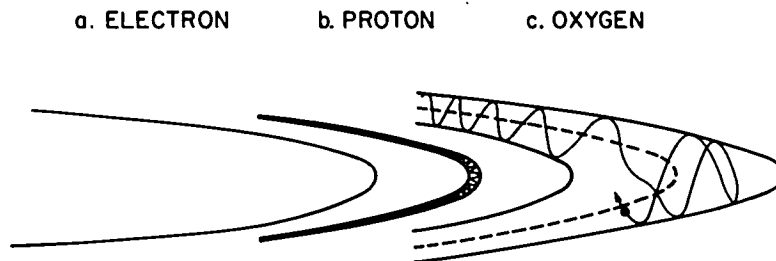


Figure 8-6. Illustrations of thermal electron and ion motions in the vicinity of a quasi-neutral sheet as exist in the Earth's magnetotail. Relative gyroradii of different particle species and realistic values of the normal field component  $B_z$  (as shown) imply that electrons are gyroscopic (a.), while ion motion can become nongyroscopic (c). (From Baker et al., 1982b).

New observational results suggest that heavy ions of ionospheric origin may, at times, dominate the composition of the plasma sheet and outer magnetosphere. The possible role of ionospheric heavy ions has been examined (Baker et al., 1982b) in the development of plasma sheet instabilities. It has been proposed in this recent work that asymmetries in the distribution of enhanced densities of  $O^+$  (of ionospheric origin) may define regions of the plasma sheet in which tearing mode growth rates are increased and the instability threshold is lowered. Qualitative estimates are made of the growth rates of the linear ion tearing mode using statistical models of the dusk-dawn and earthward-tailward distributions of  $O^+$ . The theoretical results predict maximum ion tearing growth rate in the range  $X_{GSM} = -10$  to  $-15 R_E$  and  $Y_{GSM} \sim 5 R_E$ . These values are in reasonable accord with substorm observations since many asymmetries in particle and field phenomena associated with substorms favor the interpretation that substorm onsets occur in the dusk sector of the near-Earth plasma sheet. Such results suggest that substorm initiation and development may be closely related to distribution patterns of ionospheric ions.

Another possible trigger mechanism may be found in the more earthward part of the magnetosphere. Magnetospheric electrons cannot carry upward field-aligned currents with a density of  $10^{-6} A/m^2$  or greater unless an electric potential drop of the order of 1 kV is set up along the magnetic field lines (Knight, 1973; Fridman and Lemaire, 1980; Chiu and Cornwall, 1980; Kan, 1982; Lyons, 1981). It has been suggested that it is this auroral potential structure which is responsible for accelerating auroral electrons, thus causing the variety of auroral phenomena. Akasofu (1981c) showed that field-aligned current density becomes  $\sim 10^{-6} A/m^2$  when the dissipated substorm power reaches  $\sim 10^{18}$  erg/sec. The current-carrying electrons are thus accelerated to 1 keV or greater, ionizing and exciting upper atmospheric atoms and molecules as they descend down to the E region of the ionosphere. As a result, the conductivity of the ionosphere will be increased. The increased conductivity demands, in turn, a higher field-aligned current density which requires a higher potential drop, resulting in a higher conductivity (Kan and Lee, 1980). This positive feedback mechanism is, in some interpretations, partially responsible for the explosive growth of auroral and magnetic field disturbances characteristically observed at substorm onset.

The behavior of aurora immediately prior to onset has been studied by Pellinen and Heikkila (1978b). They find a decrease in auroral brightness for a few minutes prior to the brightening of aurora at the time of onset. Further observational support for the fading of optical aurora (Baumjohann et al., 1981) and X-ray counting rates (Pellinen et al., 1982) has established this phenomenon as being intimately related to the occurrence of auroral breakup. Pellinen and Heikkila (1978a) propose that the fading phenomenon indicates the importance of inductive electric fields during substorm onset. These inductive electric fields have most recently been observed by Aggson et al. (1983).

### External Trigger

Even if the substorm onset is basically an internal process it is quite possible that an internal process is triggered by a change in the solar wind. A solar wind pressure change has been considered as a candidate for this role and indeed there is evidence that interplanetary shocks associated with storm sudden commencements often, but not always, trigger substorms (Kokubun et al., 1977).

Additional work has shown that a northward turning of the interplanetary magnetic field, or even the addition of small northward component to a stronger southward field also triggers substorms. This was first suggested by a statistical study of Caan et al. (1978) and it has been supported by further work of Caan et al. (1979) and Pellinen et al. (1982). Also a recent detailed study by Rostoker et al. (1983) has shown quite detailed correspondence between northward turnings and individual intensifications of a substorm.

## Feedback

The observations of Baumjohann et al. (1981) indicate that the ionosphere does not just play a passive role in the substorm process. It reacts to the higher conductivity caused by the particles precipitating during the substorm through the generation of polarization electric fields and production of a Cowling channel. Thus, increased dissipation of substorm energy by the ionospheric part of the substorm current wedge occurs. It may well be that there exists a positive feedback mechanism between the magnetosphere and ionosphere once the substorm process has been initiated in the tail. Such mechanisms have been described by, for example, Oguti (1971), Bostrom (1974), and Atkinson (1979). The ionospheric oxygen mechanism of Baker et al. (1982b) discussed in the Section entitled "Internal Trigger" could also be considered a feedback mechanism.

As discussed above ("Internal Trigger"), another positive feedback loop has been considered by Akasofu and Kan (1982). Starting from the hypothesis that the auroral break-up is initiated by the need to draw field-aligned currents in excess of  $10^{-6}$  A/m<sup>2</sup>, they considered that the increased particle precipitation will lead to increased ionospheric conductivity and currents. This in turn leads to the need to draw even more field-aligned current and results in a higher energy of the precipitating particles and so a positive feedback loop is set up.

## Location and Localization

Studies based on ground observations have revealed that substorms often have multiple expansion phase onsets (e.g., Clauer and McPherron, 1974; Wiens and Rostoker, 1975; Pytte et al., 1976a). Each onset is associated with a westward traveling surge, an intensification of the westward electrojet, and a Pi2 geomagnetic pulsation (Rostoker et al., 1980). Pytte et al. (1976b) examined the particles and the magnetic field observed by OGO 5 during multiple-onset substorms. It was found that the magnetic field and energetic particles in the midnight sector of the near-Earth plasma sheet showed substorm onset signatures in almost one-to-one correspondence to successive substorm onsets observed on the ground. Pytte et al. (1976b; 1978b) and Baumjohann et al. (1981) proposed that successive onsets may be repeated in the same local time sector of the magnetotail. On the other hand, Rostoker and Camidge (1971) and Wiens and Rostoker (1975) proposed that successive onsets in the magnetotail progressively shift westward.

A recent study (Nagai et al., 1983) utilized satellite data from widely spaced locations throughout the dusk-to-post-midnight sector to identify where (in local time) substorm onsets are initiated. The ground magnetic variations in a wide area were examined to determine the general characteristics of substorm activity and the timing of various phenomena. The substorm evolutions showed a variety of forms at synchronous orbit but they could be characterized by the following view: substorm activity initially starts in a longitudinally limited sector which is situated in the pre-midnight region and the effect of this activity is not seen in other regions. (For an example of how longitudinally limited the initial active region may be see the section entitled "The Substorm Current Wedge" and Figure 8-14). The activity expands westward and/or eastward during a period of successive onsets. There was an indication that the westward propagation was faster than the eastward propagation. It was suggested that the main perturbed region shifts progressively in the longitudinal direction in association with the successive onsets.

There seems to be one more notable point concerning the location of the initial onset region. The initial auroral brightening occurs typically in or south of the Harang discontinuity where also the westward traveling surge can be found (Akasofu et al., 1965; Heppner, 1972; Kisabeth and Rostoker, 1974; Baumjohann et al., 1981; Inhester et al., 1981). If one takes into account the

statistical finding that the northern part of the northward electric field region (the eastward electrojet region) is the transitional region between the central plasma sheet (cps) and boundary plasma sheet (bps) precipitation (Winningham et al., 1975), one can suggest that auroral break-ups occur on magnetic field lines connected with the transition region between cps and bps, which agrees with more direct observations by Lui and Burrows (1978).

## RELEASE

### A Model for Energy Release

#### *Description of the Model*

An optical view of what may be happening in the Earth's magnetotail is provided by astronomical observations of comets. It is observed that sections of comet tails sometimes become disconnected and drift away, resulting in a sudden loss of substantial energy and plasma to the solar wind (see Figure 8-7). Such rapid losses of energy and plasma to the on-flowing solar wind have been deduced from spacecraft measurements in Earth's magnetotail where it appears that, intermittently, a portion of the plasma sheet is severed from Earth and flows away with the solar wind (Hones, 1976, 1977, 1979). It is likely that this is a basic causal phenomenon underlying the sudden energy release aspect of a magnetospheric substorm and that the substorm's defining signatures at Earth are only some of the consequences of this rapid loss of energy and plasma from the magnetotail.

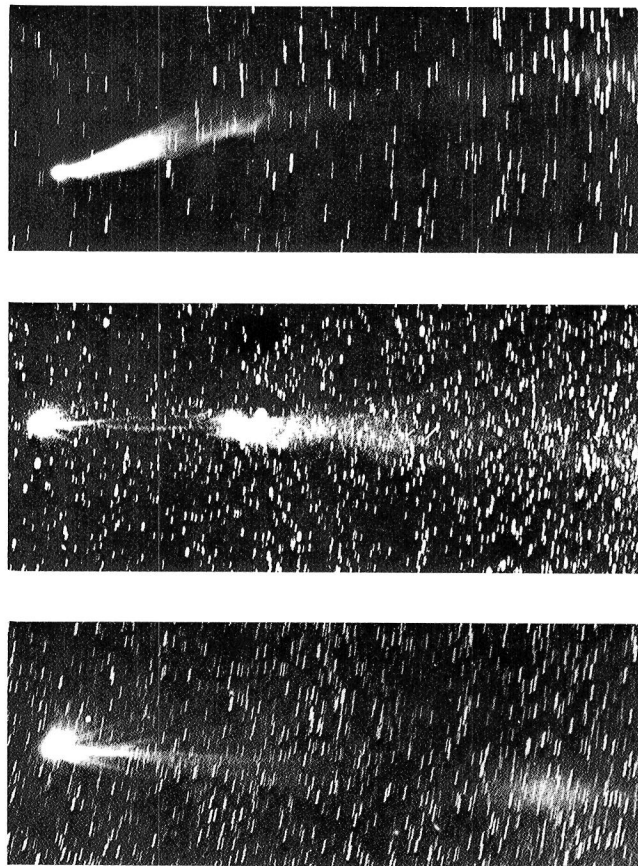


Figure 8-7. Comet Morehouse (1908 III) on September 30, October 1, and October 2, top to bottom, respectively. This sequence shows the disconnection and drifting away of the plasma tail. (Yerkes Observatory photograph).

In this section we will describe a "neutral line model of substorms" that portrays this mode of energy release from the tail and whose development has involved an interplay and blending of theoretical ideas and experimental observations that has continued over nearly two decades. The model provides remarkably good ordering of many different types of plasma and field observations within a single theoretical framework. There is also impressive agreement between this empirical neutral line model and the results of analytical theory as well as computer modeling of magnetospheric dynamics. Observational evidence for the model, as well as some points that have been raised questioning it, will be discussed.

Figure 8-8 (Hones, 1980) shows the configuration changes that the magnetotail plasma sheet is thought to undergo during a substorm. The changes depicted here constitute the central elements of the tail's role in the neutral line model. Panel 1 depicts the quiescent plasma sheet as it may exist some considerable time (e.g., an hour) before a substorm. The plasma sheet (shaded) is bounded by the field line, 5, that goes through a neutral line, N, in the very distant tail. Plasma, flowing from the lobes into the region surrounding this distant neutral line and exiting earthward along the midplane, supplies the quiescent plasma sheet and is threaded with closed field lines, 1-4, that reconnected earlier at N and that are slowly contracting earthward. Plasma ejected tailward from N is threaded with field lines that no longer have any Earth connection but are, instead, connected to the interplanetary magnetic field. The open field lines, 6 and 7, having one foot tied to Earth, are in the lobe.

Panel 2 depicts conditions at the instant of onset of the substorm's expansive phase (which is evidenced at Earth by sudden auroral brightening, occurrence of Pi2 pulsations, sharp onset of auroral zone negative bays, etc.). Before this instant, some stretching of the plasma sheet field lines had occurred that set up conditions appropriate for a "substorm neutral line," N' to form, extending for some distance across the earthward, near-midnight section of the plasma sheet, and for reconnection to start there on closed field lines. (These precursory processes constitute the "growth phase" of the substorm.) It is this beginning of reconnection in the near-Earth (i.e.,  $r \approx 10-15 R_E$ ) section of the plasma sheet that is thought to correspond to the substorm's expansive phase onset. An early consequence of the reconnection is that the long loops of the plasma sheet field lines begin to be severed, forming closed loops tailward of N' and much shortened closed field lines earthward of N'. The jetting of plasma from N' then carries the closed loops tailward and the shortened closed field lines earthward at quite high speed (hundreds of km/sec).

Panel 3, 4, 5 depict a continuation of the process of severing of plasma sheet field lines at N' as lines labeled 3, 4, and 5 successively reach N' and the number of closed loops and of closed field lines tailward and earthward, respectively, of N' increases. The reconnection of line 5, the bounding line of the plasma sheet, introduces important new consequences. The entire region of the plasma sheet tailward of N' thereafter has no magnetic connection to Earth and is free to flow tailward under the influence of the plasma pressure gradient no longer balanced by magnetic tension. This Earth-free tailward-flowing portion of the plasma sheet, threaded with closed magnetic loops, is called a "plasmoid" and is in some respects analogous to the blob of plasma leaving Comet Morehouse in Figure 8-7. Notable features of the plasmoid are that it has a magnetic neutral line of the O-type (an O-line) within it; tailward of the O-line the magnetic field has a northward z-component and earthward of the O-line the z-component is directed southward.

PLASMA SHEET CONFIGURATION CHANGES DURING A SUBSTORM

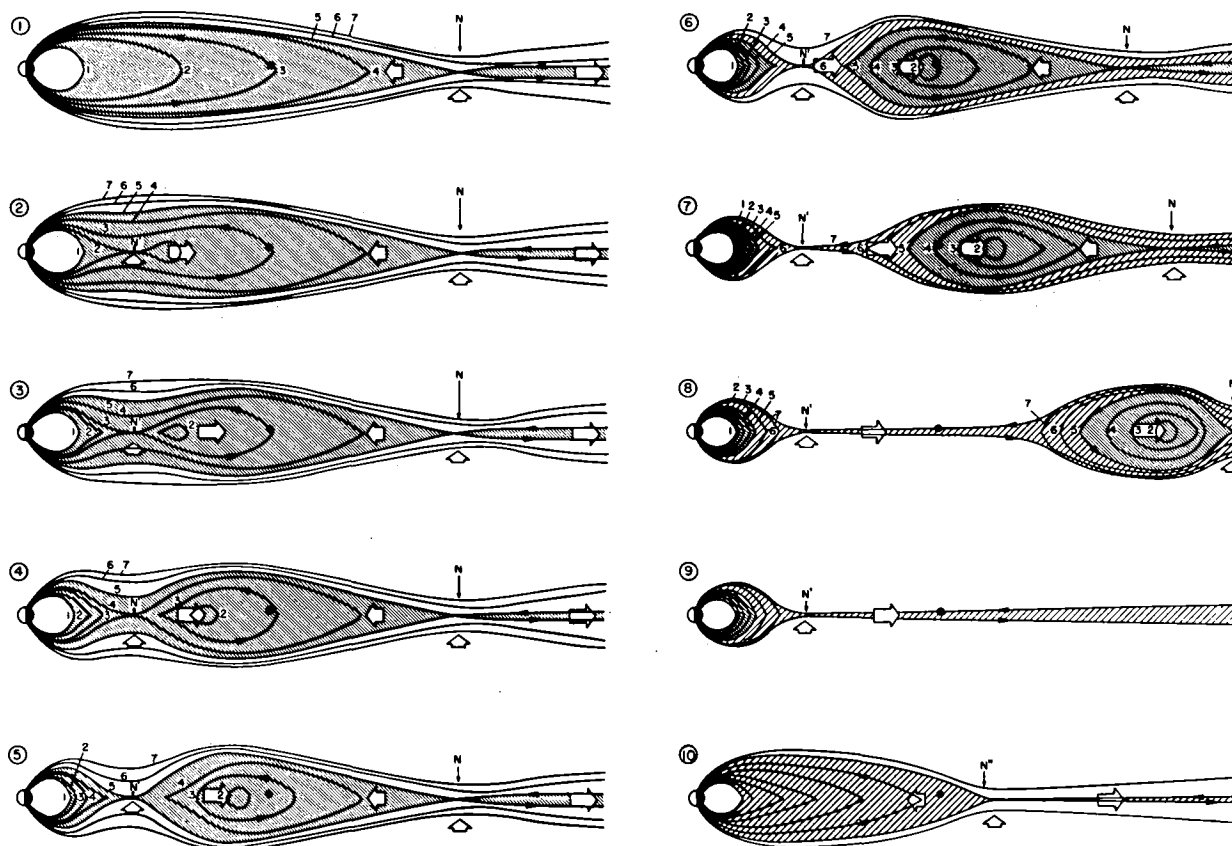


Figure 8-8. Plasma sheet configuration changes during a substorm according to the neutral line model. N is the distant pre-substorm neutral line. N' is the substorm neutral line, N'' represents the retreating neutral line at substorm recovery. A black dot represents an observing satellite at  $X \approx -30 R_E$ . White arrows depict plasma flow. (From Hones, 1980).



NOTE: In our discussion of Figure 8-8 we do not imply that the processes and magnetic configurations shown prevail across the entire width of the plasma sheet. Most probably they start near the midnight meridian, initially affecting only a small longitudinal section. The affected region then grows in width until, when the plasmoid is fully formed, its width may approach half the width of the plasma sheet (i.e., half the width of the magnetotail).

In panels 6-9 lobe field lines are reconnecting at  $N'$ . The region occupied by these reconnected lobe field lines is represented by coarse hatching in the figure. In the meantime the plasmoid is moving tailward rapidly and, by the time of panel 9, has gone beyond the border of the figure ( $\geq 60 R_E$ ). The total elapsed time between panels 2 and 9 is about 10 minutes.

The plasmoid's departure leaves a very thin plasma sheet tailward of  $N'$ . Within this thin downstream plasma sheet, plasma flows rapidly tailward and is threaded by field lines that have a southward  $z$  component and that are connected at both ends to the interplanetary field.

Panel 10 depicts what is called the recovery of the plasma sheet. This is a feature of the substorm model whose occurrence has, as yet, no theoretical explanation. Yet, observationally, this feature is one of the most predictable and one of the most impressive of the substorm sequence. It appears that the substorm neutral line,  $N'$ , after remaining in its near-Earth location for as long as a half-hour to one or two hours, suddenly moves tailward very rapidly (depicted by  $N'$  in panel 10) to some as yet unknown distance. As it does so, the plasma jetting earthward from it, threaded by newly reconnected closed field lines (with northward  $z$ -component), replenishes the plasma sheet, causing it to increase in thickness as well as length. The growing plasma sheet is shown with coarse hatching in panel 10 because it, of course, contains reconnected lobe field lines. The plasma sheet's recovery (typically occurring long -- 30-90 minutes -- after expansive phase onset) often is associated with magnetic signatures at Earth that imply that the auroral electrojet in the midnight sector is moving suddenly to an even more poleward location; this is referred to as the "poleward leap" of the electrojet.

In this description of the model we have said little thus far about the situation earthward of the substorm neutral line,  $N'$ . As Figure 8-8 shows, the newly reconnected closed field lines, first those of the plasma sheet and later those of the lobe, contract earthward, carrying plasma inward and presumably projecting some of it into the high latitude (auroral) ionosphere. Thus, the substorm signatures seen at Earth are caused almost entirely by the processes earthward of  $N'$  which are, in a sense, by-products of the principal action of the magnetotail, i.e., that of creating and releasing the plasmoid whereby the magnetosphere divests itself of excess plasma and energy that it can no longer restrain.

### *Observational Bases for the Model*

The neutral line model derives its major support from satellite observations of magnetic fields, plasmas, and energetic particles made with satellites in the region of the magnetosphere extending from the geosynchronous orbit (at  $6.6 R_E$ ) out to  $\sim 35 R_E$  in the tail. Alternatively, it can be said that the model has been developed from such observations. The present discussion will treat the observations under four broad topical categories: Plasma Sheet Thinning;  $B_z$  Polarity; Plasma Flow; and Open Field Line Identification.

### 1. Plasma Sheet Thinning and Recovery

One consequence of reconnection that is evident in Figure 8-8 is that a satellite that is in the plasma sheet at  $x < -15 R_E$ , i.e., tailward of the substorm neutral line, and more than 1 or 2  $R_E$  above or below the neutral sheet (i.e., the midplane of the figure), ought to witness a disappearance of the plasma sheet soon after substorm onset (e.g., panel 7 or 8 in the figure). It should also witness a reappearance of the plasma sheet as it thickens (panel 10). This process, now referred to as plasma sheet thinning and plasma sheet recovery (or thickening) has, indeed, been observed and widely studied (e.g., Hones et al., 1967; Lui et al., 1975). In the vicinity of  $r \sim 18 R_E$  where the process was first observed, the plasma sheet thinning coincides closely with the onset of the substorm expansive phase. The subsequent thickening then occurs as the substorm-associated auroral zone negative bays subside (often 30 minutes to an hour or more after thinning). At somewhat greater distances from the Earth the thinning noticeably lags expansive phase onset, occurring typically 5 to 10 minutes after the onset. These processes change character at  $x \approx -15 R_E$  (Hones et al., 1973). Earthward of this location the plasma sheet thickens, rather than thins at expansive phase onset.

Although the plasma sheet thinning and thickening are suggestive of magnetic reconnection in the manner shown in Figure 8-8, they are not a unique signature of the process and they have been interpreted in terms of at least one other model (Chao et al., 1977).

### 2. $B_z$ Polarity

A second consequence of magnetic reconnection that is apparent in Figure 8-8 and that is unique and potentially observable in the tail, is the occurrence of southward  $B_z$  in the neutral sheet tailward of the x-line, N'. Historically the observation of southward  $B_z$  in clear association with substorms has proven difficult and the reason for this can be seen in Figure 8-8. It is evident there that southward  $B_z$  will occur in the trailing (earthward of the O-line) portion of the departing plasmoid and in the thin downstream plasma sheet after the plasmoid departure. The plasmoid is relatively thick but moves tailward very rapidly (500-1000 km/sec), so a satellite located as far as several Earth radii above or below the midplane can obscure the southward  $B_z$  in the plasmoid's trailing portion, but it can do so for only a few minutes. The downstream plasma sheet is very thin (perhaps a few thousand kilometers) so a satellite has little chance of being within it and of sensing the southward  $B_z$ . Furthermore, the downstream plasma sheet is very dynamic and wavy (Fairfield et al., 1981b) and can display northward and southward  $B_z$  components due to these factors that can obscure the  $B_z$  variations due to reconnection.

Despite the obvious difficulty of these observations, Nishida and Nagayama (1973) were able to show in a study of many substorms, that in the plasma sheet region,  $-25 R_E > x > -35 R_E$  and  $|Y| < 10 R_E$ ,  $B_z$  became negative almost without fail near the onsets of strong substorms. They concluded that a neutral line is formed somewhere between  $x = -10$  and  $-30 R_E$  extending across the sector  $|Y| \lesssim 15 R_E$  within about 10 minutes of the expansive phase onset. They concluded, further, that the neutral line continues to exist in its near-Earth location for about 1 hour and then disappears almost as quickly as it is formed.

### 3. Plasma Flow

A third consequence of reconnection, apparent in Figure 8-8, is the rapid tailward flow of plasma from the reconnection region. This feature, also, is not unique to a reconnection model and has been interpreted in terms of other models. But the measurements of plasma bulk flow have provided some of the strongest support for the neutral line model, especially when they have been

combined with determinations of  $B_z$  polarity. This follows because, as Figure 8-8 shows, tailward flow of plasma threaded with southward field is expected tailward of the substorm neutral line while earthward plasma flow threaded with northward field is expected earthward of it.

The first combined measurements of plasma flow and magnetic field (Hones et al., 1976) revealed a sequence of phenomena that has subsequently been observed many times and that is understood in the context of Figure 8-8. Precisely (+1 minute) at expansive phase onset a satellite at  $X \sim -30 R_E$  detects fast (300-500 km/sec) tailward flow of plasma. Initially  $B_z$  is northward in this tailward flowing plasma but the field turns southward after  $\sim 1$ -2 minutes. After another few minutes the plasma density decreases markedly sometimes suffering a "dropout" below the level of detectability. This situation then prevails for 30 minutes or more, during which the plasma sheet may reappear for a few-minute interval and then disappear again. Within this thinned plasma sheet, plasma flows tailward and  $B_z$  is often southward and sometimes is the major component of  $\vec{B}$ . As the substorm's expansive phase begins to subside (as determined from geomagnetic records from auroral latitudes) the plasma sheet reappears permanently and features earthward flowing plasma threaded with northward magnetic fields.

#### 4. Open Field Line Identification

The observations we have described so far support the idea of neutral line formation in the near tail at expansive phase onset and its tailward retreat during substorm recovery. We next discuss the evidence for perhaps the most prominent feature depicted in Figure 8-8, i.e., the formation and release of a plasmoid during the substorm sequence. For this, one may use a measurement of energetic ( $>200$  keV) electrons that allow differentiation between open and closed field lines. These electrons travel distances comparable to the magnetotail's dimensions in about a second and it has been shown that when they are observed on open field lines they possess a unidirectional anisotropy indicative of tailward streaming (Baker and Stone, 1976). This means of identification has been used by Hones (1977, 1979) and by Bieber et al. (1982) in analyses of a number of substorms using data from a satellite (IMP 8) at  $X \sim -30 R_E$ . An essential result of those analyses is the recognition that the plasma found flowing tailward for the first few minutes after expansive phase onset (i.e., before the plasma dropout) is threaded by closed magnetic field lines (i.e., the  $>200$  keV electron flux is quasi-isotropic); open field lines are encountered only when the boundary of the thinning plasma sheet reaches the satellite. Equally important, it is noted that for the initial minute or so of tailward flow  $B_z$  remains northward before turning southward. These observations suggest (as in Figure 8-8) that reconnection starts in the closed field lines of the pre-substorm plasma sheet and that a satellite at  $\sim 30 R_E$  is initially tailward of the O-line of the resulting configuration of closed magnetic loops (the plasmoid). The crossing of the departing plasmoid's trailing surface is marked by the  $>200$  keV electrons' transition from quasi-isotropic to tailward streaming.

#### Some Statistical Results

Bieber and Stone (1980) searched several years of  $>200$  keV electron data from Imp 7 and Imp 8 for occurrences of tailward streaming of those particles. They found about 50 such events and observed that they tended to occur near the onset of substorms. Later, Bieber et al. (1982) found that concurrent plasma data were available for 16 of those events, and they performed a superposed epoch analysis, for the 16 events, of the plasma density, the plasma bulk flow components,  $V_x$  and  $V_y$ , the magnetic field z-component, and the AE index. The zero epoch time,  $T=0$ , was determined by the first occurrence of intense tailward streaming of  $>200$  keV electrons in each event. Figure 8-9 shows the results of that analysis. Most of the features implied by Figure 8-8 are seen in Figure 8-9. Onset of electron streaming (indicative of satellite passage onto open field lines) is preceded by  $\sim 5$  to 10 minutes of rapidly increasing AE, fast tailward plasma flow (negative  $V_x$ ), and southward  $B_z$ . Minutes later the plasma density drops to very low values

and  $B_z$  becomes 0, indicative of lobe environment, signifying the departure of the plasmoid. At about  $T=+30$  minutes AE decreases rapidly,  $B_z$  becomes strongly positive,  $V_x$  becomes positive, and the density begins to rise toward its pre-substorm level. These last points are the signatures of plasma sheet recovery (panel 10 of Figure 8-8) that results when the substorm neutral line retreats suddenly and rapidly tailward.

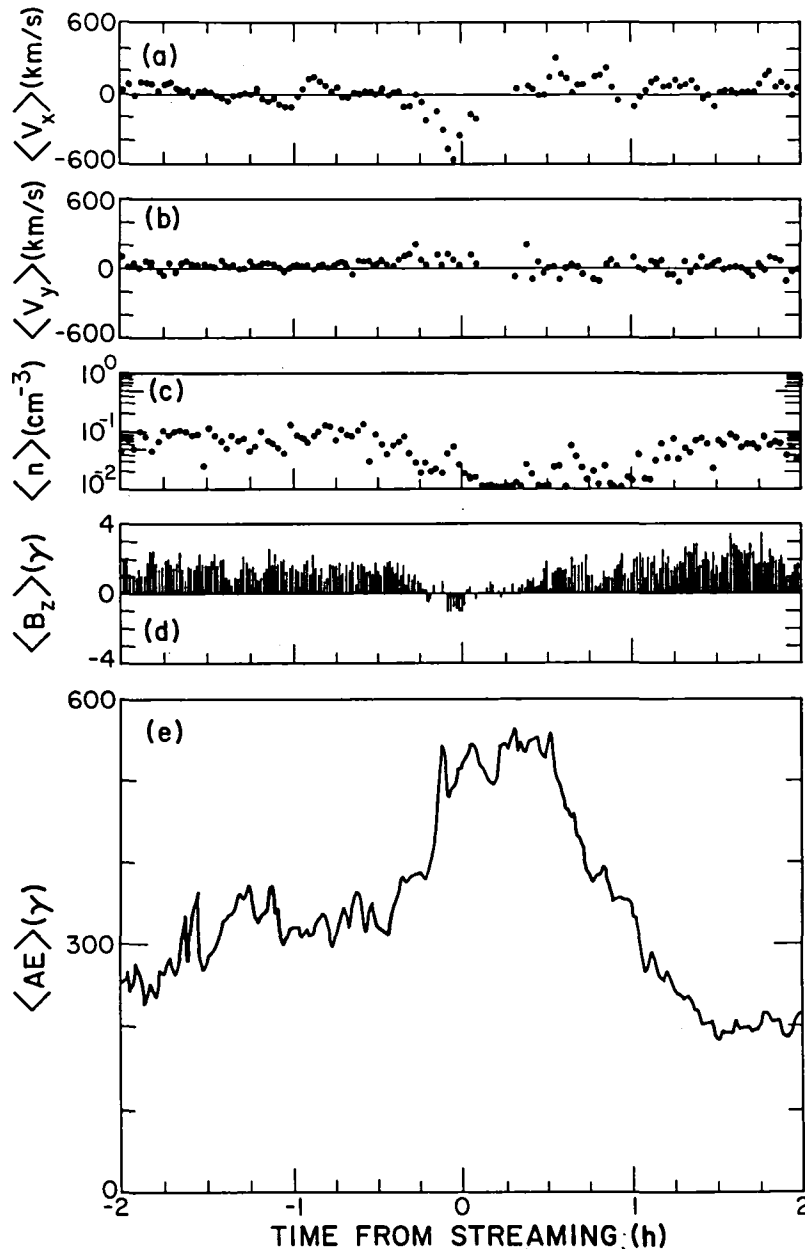


Figure 8-9. Superposed epoch analysis of (a)  $X_{GSE}$ -component of the plasma flow velocity; (b)  $Y_{GSE}$ -component of plasma flow velocity; (c) plasma density (plasma parameters are 2-minute averages for protons of energy 84 eV to 15 keV); (d)  $Z_{GSM}$ -component of magnetic field (1-minute averages); (e) AE index (1-minute averages). Plotted values are the median over 16 events. Zero epoch time is determined by the first observation of intense tailward streaming of  $>200$  keV electrons. (From Bieber et al., 1982).

### *Alternatives to the Neutral Line Unloading Model*

The reconnection signatures discussed here have been shown by numerous statistical and event studies to occur rather persistently during substorms, but they are by no means universal. A complete theory of substorms must explain both the events in which the signatures are present and those in which some of all of them are absent. Many investigators favor the interpretation that reconnection in the near-Earth plasma sheet is, indeed, the underlying feature of substorms but that the total process is more complex than the present model depicts. Furthermore, they believe that observations of substorm-related phenomena that depart from the predictions of this neutral line model may provide important information on aspects of reconnection which are not completely understood at present. Alternatively, the absence of reconnection signatures during some substorms had led other researchers to reject the reconnection theory altogether and to seek other underlying models for the rapid substorm energy dissipation process.

In a series of papers, Lui et al. (1976; 1977a, b) reexamined magnetic field and plasma data which had previously been taken as supportive of the neutral line model. They reached conclusions that were at odds with the neutral line model, at least in its simplest form. These results led to an alternate suggestion for the cause of substorm-related magnetotail dynamics, viz, the propagation of a rarefaction wave down the tail following an instability in the near-Earth plasma sheet (Chao et al., 1977). Further objections to the standard model were raised by Akasofu et al. (1978) who argued that in its (then) 2-dimensional form it lacked conviction because it could not explain certain 3-dimensional variations of the tail magnetic field that occur systematically during substorms.

More recently, several researchers have examined three-dimensional measurements of plasma and energetic ions in the plasma sheet boundary and find counterstreaming beams of ions at times when moment calculations have shown unidirectional flow of plasma. It has been suggested that plasma flows, derived from moment calculations (that do not portray these beams), may be misleading where moments have been used in support of the neutral line model. While it is not clear that any substantially different conclusions regarding the basic validity of the neutral line model would result from a more complete representation of these boundary distribution functions, they warrant further examination because of a more complete understanding of the model that will result.

It is clear from all of the above criticism that continued examination and elaboration of the neutral line model is very helpful in understanding the multifarious aspects of the sudden energy-release component of substorm dynamics. For example, Speiser and Schindler (1981) examined observational evidence (using neutral sheet magnetic field data) and generally found results consistent with near-Earth tearing mode development and inconsistent with the rarefaction wave model (as well as other suggested alternatives to near-Earth reconnection). Similarly, a more thorough elaboration of the model and examination of the data in this light (Hones and Schindler, 1979) appears to answer many of the criticisms of Lui et al. and, in fact, shows that many of the discrepancies that Lui et al. found with the model are quite consistent with the predictions of the more detailed treatment. In addition, the three-dimensional magnetic variations during substorms, discussed by Akasofu et al. (1978), have been shown by observational results (Hones et al., 1982) and by 3-D computer modeling of magnetotail reconnection (Birn and Hones, 1981) to be a predictable consequence of the reconnection process when its three-dimensional aspects are taken into account.

### **Impulsive Plasma Injections**

The high altitude phenomenon which is most directly linked with polar substorm currents and auroras at ionospheric heights is the impulsive injection of a hot plasma (multi-keV) and

energetic particle population into a region of space where such a plasma would not otherwise be present. Plasma having an energy density, which (assuming strong pitch angle diffusion) is sufficient to produce all but the brightest auroras, routinely appears near synchronous orbit in association with substorm magnetic disturbances. Such a plasma is not present at that location during magnetically quiet periods. The origin and mechanism of injection of this plasma are therefore issues of some importance in understanding magnetic substorms.

Early observations of particles having energies below those in the radiation belts were hampered by the lack of a technology capable of detecting individual electrons or ions. Using data from a Faraday cup instrument having only four energy windows (40 eV - 2 keV), on the eccentrically orbiting OGO 1 and 3 spacecraft, Vasyliunas (1968) was able to deduce most of the important features of the plasma sheet inner edge. In particular, he identified the shift of the characteristic electron energy to lower values without significant change in the density as one passes radially inward through the boundary. Moreover, Vasyliunas argued that electron increases resembling a reversal of the normal inbound pass were produced in one case by transient earthward motions of the plasma sheet from about  $11 R_E$  to  $6 R_E$ . This phenomenon was clearly associated with the ground signatures of a substorm.

When appropriate instruments were flown aboard geosynchronous satellites traveling nearly tangent to the inner plasma sheet, new perspective was gained on the substorm effects there (Arnoldy and Chan, 1969; DeForest and McIlwain, 1971). Abrupt appearances of plasma and energetic particles occurred on the nightside of the orbits. Such "injections" evidently create longitudinally localized "clouds" of plasma whose energy-dependent drift lead to elegant dispersion patterns. Although a persistent low energy feature was observed whose position was simply dependent upon the general level of activity (Mauk and McIlwain, 1974), it became evident that a full description of this region of space would require a model considerably more complex than a simple advancing boundary (McIlwain, 1974; Konradi et al., 1975; Mauk and Meng, 1983a).

### *Substorm Particle Injections*

Substorm particle injections observed near the geosynchronous regions have the following characteristics:

- The energy dispersion of the quasi-Maxwellian, quasi-isotropic plasmas (0 to 100 keV ions and electrons) is extremely complex and variable. Many different types of dispersion events can be identified and many of these have dispersion sense symmetry properties which are distinct from those expected from ordinary convection (DeForest and McIlwain, 1971; McIlwain, 1972; Mauk and Meng, 1983a).
- Some injections appear to have leading edges (on an energy-time spectrogram) which show virtually no energy dispersion simultaneously in both ions and electrons to very high (>100 keV) energies (Arnoldy and Chan, 1969; Bogott and Mozer, 1973; Baker et al., 1978, 1979a).
- Dispersionless injections can form inside the geostationary orbit where the field lines are clearly closed and dipolar in configuration both before and after the injection. The evidence suggests that dispersionless injections do not necessarily appear at any magnetic topological transition region such as the region between open and closed field lines.
- Dispersion pattern results from older substorm injections can appear, on a greyscale spectrogram, to pass through and be virtually undisturbed by more recently generated dispersion patterns (McIlwain, 1974; Mauk and Meng, 1983a). On high resolution displays, even the fine energy structure is preserved in the old plasmas. The transition across a dispersion contour between the "old" undisturbed plasmas and the "new" plasmas can be extremely sharp.

All four of the points described above, but most particularly the fourth point, prompted McIlwain (1974) to suggest the "injection boundary" model (see also Konradi et al., 1975; Mauk and Meng, 1983a, b). The injection boundary model suggests that near the initiation time of a substorm expansion phase, particles are energized and/or transported very quickly (on time scales much faster than ordinary convection times) so that all energies of the distribution fill at one time an extended region which lies tailward of a sharply defined, nightside, spiral-like injection boundary.

In order to explain the many classes and characteristics of the energy dispersion features observed, and their scaling characteristics, the injection boundaries must have a configuration which is quite distinct from the well known convection boundaries, or Alfvén layers. The present best guess for that configuration (Konradi et al., 1975; Mauk and Meng, 1983a, b) is that in the dusk to near-midnight quadrant, the boundary (at the initiation of the injection only), spirals toward the Earth as one moves toward midnight from the dusk side. In the near-midnight to dawn quadrant the injection boundary spirals away from the Earth as one moves downward from near midnight.

A key strength of the "injection boundary" model is that it shows that the many classes of observed dispersion events can be understood as a single, common sequence of events. Its weakness is that it is phenomenological in nature, and that it does not tell us what the mechanisms are that give rise to the transport of particles from the tail regions. Attempts have been made to use time stationary and perturbed convection to account for the transport (e.g., Chen, 1970; Kaye and Kivelson, 1979; Kivelson et al., 1980). However, only the limited features of injections have been used to support such models, and it appears that virtually all electron-ion dispersion features observed in the geostationary orbit, even during quiet times, have some characteristics which are inconsistent with the simple convection models that have been proposed (Mauk and Meng, 1983a, b). It may be that by the addition of further complexities the convection models might explain more features than they presently do. However, it appears at present that the near geostationary regions ( $r < 8 R_E$ ) are not populated by ordinary global convection, either perturbed or time stationary (Mauk and Meng, 1983b). Thus, the importance of transient effects can never be ignored.

### *Multipoint Injection Observations*

Single-point observations leave some room for doubt as to the relative role of intrinsic and convective temporal changes, and conflicting views arose in connection with the nature of synchronous orbit plasma injection events (McIlwain, 1974; Kivelson et al., 1980). The issue is that the injection could result either from a time-varying, spatially localized heating process with subsequent dispersal in an unchanging flow field, or from a time-varying flow of plasma with constant properties, or from some combination of the two.

Multipoint observations which began to be available in the mid-70s have, to some extent, resolved these ambiguities by unequivocally demonstrating the earthward motion of a boundary between soft inner plasma sheet electrons and hot central plasma sheet electrons, in association with substorms. Barfield et al. (1977) were the first to report such a motion. Using data from ATS-6 and Explorer-45, they found a time scale for the motion over a  $1.2 R_E$  baseline of 11 minutes, implying that plasma was moving earthward in a region from which it was previously excluded. Moore et al. (1981) studied several similar events using data from ATS-6 and SCATHA (P78-2) with particular attention to the abrupt and dispersionless nature of the leading edges of many events. They found that events which were abrupt at both spacecraft traveled  $1 R_E$  in as little as one minute, implying average speeds up to 100 km/s and boundary thickness  $\sim 0.1 R_E$ . They further argued, on the basis of electron energy spectral changes that the moving plasma boundary which they referred to as the "injection front," was the precipitation-flow boundary

described by Kennel (1969). The events studied by Moore et al. (1981) were closely associated with large (factor 2-3) equally abrupt increases of local magnetic field strength and with rotations to a more dipolar orientation. On the basis of this fact, the agent of injection was identified as the earthward propagating compression wave previously observed by Russell and McPherron (1973). Such a compression wave will steepen and displace earthward any quiescent structures present in the plasma through which it propagates and thus qualitatively account for a sharp earthward moving boundary and for the displacement of plasma into regions where it would ordinarily be excluded.

These studies, while emphasizing the role of rapid boundary motion, have not indicated a total lack of local heating or acceleration. On the contrary, Moore et al. (1981) argue that multiple injections in a given evening had cumulative effects on the hot plasma being swept over the spacecrafts. Though the quasi-Maxwellian (multi-keV temperature) hot electron distribution was relatively unchanged in successive injections, a nonthermal power law tail of the distribution was enhanced at each injection. They speculated that an appreciable amount of each compression wave's energy was being dissipated in the inner plasma sheet. Evidently both views are correct: transient flow changes and boundary motion dominate the plasma effects while local acceleration dominates the energetic particle changes.

It should be noted that the compression wave observed by Russell and McPherron (1973), and hypothesized as an agent of injection, contains a dawn-to-dusk directed current sheet and has an associated dawn-to-dusk directed induced electric field ( $\sim 10$  mV/m) which is consistent with recent electric field observations (Pedersen et al., 1978; Shepherd et al., 1980; Aggson et al., 1983). The motion of such a wave corresponds directly to the collapse of stretched magnetic field lines to a less stretched configuration (cf., Baker et al., 1982a), the collapse being communicated to more earthward locations by the propagating wave. An extremely interesting aspect of such a wave is that it propagates into a region of decreasing phase speed. This is due to the increase in plasma density and decrease in ion mean energy associated with the plasmasphere, which is often appreciable near synchronous orbit. In such circumstances, the wave will surely steepen and may break, forming a shock, or be reflected. In either case, it should be unable to propagate very deeply into the plasmasphere, and there should be a well defined earthward limit of injection effects.

#### *Relationship of Injection Fronts and Injection Boundaries*

The compression wave hypothesis addresses the transport of tail plasmas, but it does not explain the dispersal features that are observed. Both transport and dispersion may be accommodated by bringing together the compression wave model and the injection boundary model discussed earlier. This can be accomplished by hypothesizing that the propagating plasma sheet inner edge (injection front) has an earthward displacement which maximizes in the midnight sector. This earthward displacement may or may not correspond to the earthward-most propagation of the compression wave. The key to the validity of this combined model would be a demonstration that the near-midnight bulge configuration, required of the injection boundary, results from the compression wave displacement in a natural way. A schematic illustration of the required behavior is shown in Figure 8-10.



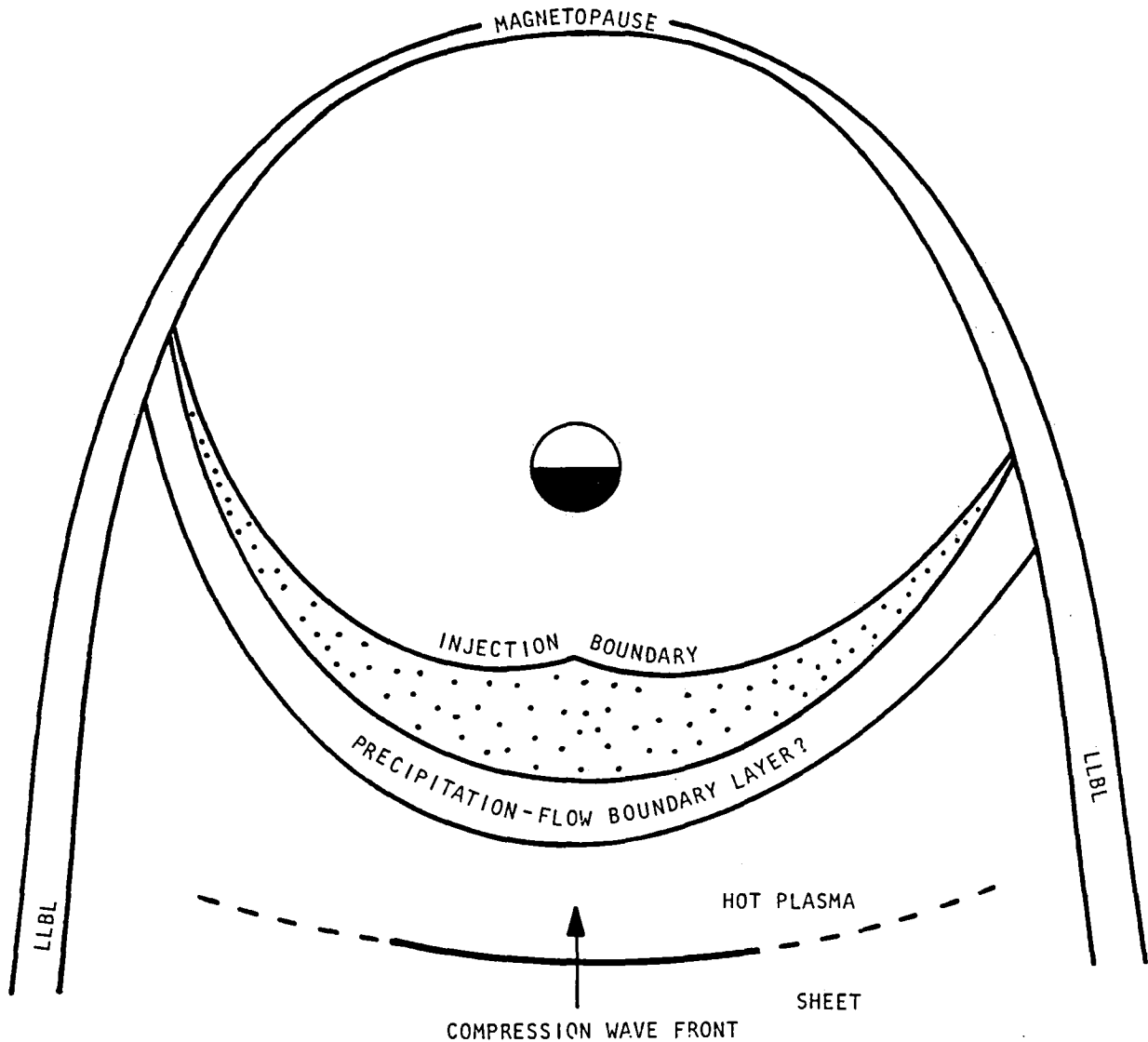


Figure 8-10. Equatorial plane schematic of an hypothesis concerning the formation of an injection boundary near synchronous orbit. A compression wave impulse propagates earthward from the magnetotail through the pre-existing inner edge of the plasma sheet, producing an earthward moving injection front with maximum net displacement in a sector near midnight.

### *Consequences of Injections*

One consequence of the injection processes described above is contribution to the formation of the ring current distributions. The ring current particles reside in the inner magnetospheric regions ( $r \sim 2-6 R_E$ ) and they give rise to depression in low latitude ground magnetic measurements (characterized by the Dst index). Evidence has been given that directly driven, non-impulsive, processes contribute to the formation of the ring current (see Section III). Such contributions are likely to take the form of enhanced global convection (Harel et al., 1981). However, as discussed earlier, the particle features observed in the geosynchronous regions are not consistent with global convective transport of particles from the outer magnetospheric region. Impulsive injection with characteristics distinct from global convection appear to dominate the earthward transport of particles near the geosynchronous orbit. These impulsively transported particles must play a role in the ring current formation, but the relative importance of the impulsive and driven processes remains to be determined. One role that the driven or convective processes may play is to transport further earthward the particles which have been impulsively transported to the geosynchronous regions. It has been shown by Lyons and Williams (1980) that distributions measured in the very innermost region ( $r \sim 3 R_E$ ) can be accounted for by adiabatically displacing particles from the regions just earthward of the geosynchronous region ( $r \sim 5 R_E$ ). This displacement could take place by means of global convective processes. The impulsive injection would then form just one leg of the processes responsible for the ring current formation.

A second consequence of impulsive substorm injections is to provide a particle reservoir which gives rise to the diffuse aurora. Meng et al. (1979) have shown that precipitated electron distributions measured over diffuse aurora match the trapped distributions measured at the conjugate geosynchronous region. From the energy dispersion of the combined ion-electron features it is clear that the trapped distributions had been impulsively injected.

### *Implications of Injections*

The observation of clearly defined impulsive injection events apparently originating within the magnetosphere virtually guarantees that there is some energy storage and instability mechanism operative in the magnetospheric tail. At intermediate levels of activity the individual substorms are separated by a time comparable to or longer than plasma loss and dispersion times, so the events are obviously discrete. In times of great activity, individual injections occur with a separation which is shorter than the time for injected plasma to disperse, so that the effects of individual injections are much more difficult to discern and represent smaller perturbations in the global energy flow. Overall, there appears to be more variability in the frequency with which substorms occur than in the magnitude of individual substorms. However, the situation becomes so complex in very active times that one cannot preclude a qualitative change in the mode of energy dissipation as one progresses from a series of fairly frequent substorms to a full fledged geomagnetic storm producing significant ring current (see Section II above).

### **Energetic Particle Acceleration**

Just as for hot plasmas (Section entitled "Impulsive Plasma Injections" above), it is now generally recognized that during the course of magnetospheric substorms substantial energetic particle acceleration occurs within the near-Earth part of the terrestrial magnetotail. Subsequent to the onset of substorm expansion phases, it is essentially always observed that energetic particles ( $E \geq 20$  keV) appear at the synchronous altitude of  $6.6 R_E$  (Arnoldy and Chan, 1969; Lezniak and Winckler, 1970; Bogott and Mozer, 1973; Baker et al., 1978). Considerable evidence, as reviewed in the previous section, has been presented to suggest that plasma heating and energetic particle production often occur within a geocentric radial distance of  $15-20 R_E$ .

A variety of mechanisms contribute to the acceleration of ions and electrons to the moderate (<100 keV) energies emphasized in the previous section. The details of many of these mechanisms are covered, for example, in the reports of Solar Terrestrial Physics Working Groups 2 and 7 and will simply be listed here:

- Parallel electric fields above the auroral oval accelerate particles to 1-10 keV
- Electrostatic ion cyclotron waves probably produce ion conics over the auroral zone
- Betatron acceleration of particles convecting from the plasma sheet to  $6.6 R_E$  (see "Impulsive Plasma Injections" above) can account for a part of the low-energy component of particle injections
- Acceleration of particles in the distant ( $\sim 100 R_E$ ) neutral sheet may account for the characteristics of ion beams in the plasma sheet boundary layer
- Shock acceleration in substorm compression waves may occur (cf. Section "Impulsive Plasma Injections")

The energization mechanisms described above do not appear capable of producing the high energy (>100 keV) component of substorm-related particle enhancements. Such particles could, however, be rapidly accelerated in the parallel electric field which exists along a near-Earth neutral line. Note that induction effects related to dynamic reconnection can raise the total potential drop along the neutral line far above its expected steady-state value (Baker et al., 1982a). Moreover, acceleration processes at the neutral line could explain the electron heating pulse observed during some reconnection events (Bieber et al., 1982). Such heating is reminiscent of the laboratory reconnection results of Stenzel et al. (1982) (cf. Working Group 1 report).

Enhancements of the fluxes of particles having energies of several hundred keV are commonly observed at synchronous orbit and in the magnetotail during geomagnetically active periods. In many cases these flux enhancements have been shown to be closely associated with individual magnetospheric substorms. Using energetic particle data from synchronous orbiting satellites and from satellites in the magnetotail, it is found that many features of the timing of particle enhancements relative to substorm onsets and recoveries (derived from ground magnetic records) and relative to plasma sheet thinnings and recoveries (measured with plasma probes on Earth-orbiting satellites) can be understood in terms of the neutral line model of substorms (Section entitled "A Model for Energy Release") in which the particles are impulsively accelerated during a brief period at substorm onset.

In approximately 10-20% of substorms, >0.3 MeV ions appear throughout the magnetosphere and its environs in close association with expansion phase onset (Belian et al., 1978). A comprehensive model for the morphology of energetic ion enhancements is illustrated in Figure 8-11 (Baker et al., 1979b). This model suggests that after acceleration at the X-line, the ions stream both sunward and tailward. Those reaching the synchronous orbit region are transported westward around Earth via curvature and grad-B drifts and may be identifiable as distinct particle bunches ("drift echoes") through several (as many as 5) circulations.

The tailward-streaming ions appear as "impulsive bursts". The inverse velocity dispersion (i.e., observation of slower particles before faster ones) exhibited by these bursts is supportive of their hypothesized origin at a magnetic X-line (Sarris and Axford, 1979). As suggested by the inset at the bottom of Figure 8-11, a spacecraft in the thinning plasma sheet successively samples field lines that have reconnected more and more recently at the X-line. These field lines contain ion distributions that are less depleted at the high energy end of the spectrum by escape

of the faster particles. Finally, just as the spacecraft enters the lobe, it samples preferentially the fastest ions streaming along field lines connected directly to the X-line source. Concurrent plasma observations confirm that impulsive bursts do indeed occur right at times of plasma dropouts (Belian et al., 1981).

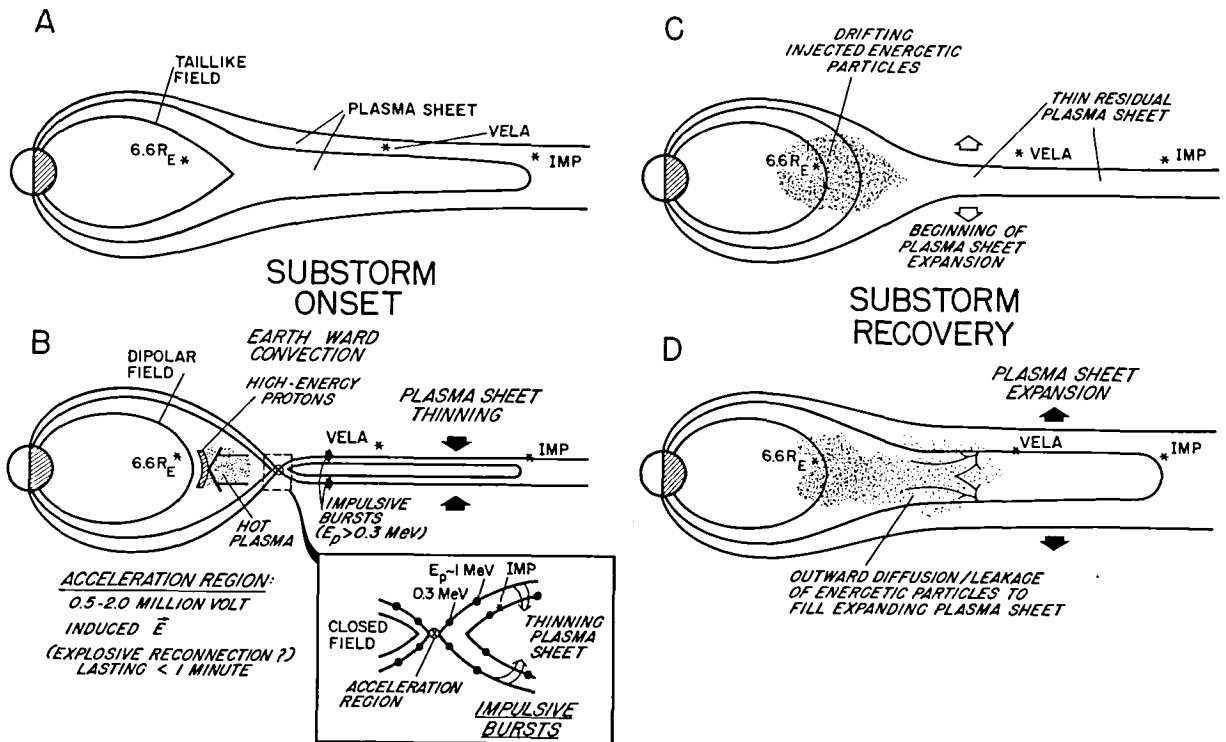


Figure 8-11. Schematic depicting the sequence of energetic particle events predicted by the model of Baker et al. (1979). (a) The inner magnetosphere just prior to substorm onset showing the buildup of stress evidenced by the tail-like field. (b) The magnetosphere just after onset showing a dipolar field configuration and the accelerated proton bunches streaming sunward toward the trapped radiation zones and antisunward along the thinning plasma sheet. (c) Conditions just prior to substorm recovery and the beginning of the plasma sheet expansion. (d) Expansion of the plasma sheet and the subsequent filling of the expanding sheet with energetic protons diffusing out of the trapped region.

As a final component of this picture, the more commonly observed non-impulsive (rapid-rise, slow-decay) plasma sheet ion enhancements are attributed in this model to envelopment of the observing satellite by the recovering (i.e., expanding) plasma sheet, into which have leaked ions previously injected into the outer radiation zone. The subsequent decay is explained by a combination of plasma sheet expansion, adiabatic cooling, and escape mechanisms. Escape from the magnetotail could in turn account for the appearance of energetic proton bursts in the magnetosheath and upstream region.

Although the number densities of energetic particles are relatively small compared to plasma number densities, the energy density in this component can be reasonably large throughout the outer trapping region and the magnetotail. Furthermore, once such particles are produced, they make excellent diagnostic tools for establishing substorm timing (e.g., Belian et al., 1981), for examining field line topology (see section on "A Model for Energy Release"), and for remotely probing plasma boundary motions (e.g., Baker et al., 1982a).

### **The Substorm Current Wedge**

A significant part of the energy stored in the magnetotail (cf., Section III "Storage") is dissipated through the ionospheric part of the substorm current wedge. This substorm current wedge is set up by the sudden disruption of an azimuthally confined section of the enhanced cross-tail current and its diversion to the auroral ionosphere via field-aligned currents (cf., Figure 8-12 and see McPherron et al., 1973, and Bostrom, 1974). While the existence of the current wedge was known for a long time (cf., Akasofu and Meng, 1969) the physical mechanism leading to its formation was unclear until recently. Strong support for the neutral line model of substorms was obtained when three-dimensional MHD simulations of magnetotail reconnection (Birn and Hones, 1981; Sato et al., 1983) showed that a pair of oppositely directed field-aligned currents are an inherent picture of the neutral line model (see Section IV "A Model of Energy Release").

Unfortunately, due to the paucity of satellites in the magnetotail not much more can be said on the magnetospheric part of the substorm current wedge than that it is azimuthally confined, has a geometry as shown in Figure 8-12 and expands eastward and (especially) westward during the course of the expansion phase (see also Nagai, 1982). However, the near-Earth part of the current wedge can be studied by ground-based observations. Furthermore, knowledge of the configuration and dynamical development of the near-Earth field-aligned currents and their ionospheric closure current may allow one to draw conclusions on the configuration and dynamical development of the whole substorm current wedge. Moreover, it is of vital importance to determine the nature of the ionospheric closure current in order to decide whether the auroral ionosphere is just a passive medium where substorm energy is dissipated or if it reacts and possibly affects the substorm development (cf., Section on "Feedback"). In the following we will illustrate what is known about the near-Earth part of the substorm current wedge by summarizing two recent case studies of substorm current systems.

Baumjohann et al. (1981) have analyzed two-dimensional ground magnetic and ionospheric electric field observations during these successive local auroral break-ups around magnetic midnight which seems to constitute the auroral signatures of a multiple onset substorm. Simultaneous two-dimensional observations of the magnetic, electric and auroral signatures made it possible to construct a fairly sophisticated model of the three-dimensional current flow in the region covered by the active aurora. Each of the auroral break-ups was associated with a sudden change in direction of the electric field from northward to southward and the appearance of a strong westward equivalent current. These observations are consistent with the generation of an incomplete Cowling channel in the highly conductive active region (e.g., Bostrom, 1964, 1974; Coroniti and Kennel, 1972). Apparently the precipitating electrons create a local region where the conductivity is appreciably higher than in the surrounding region. The westward component of

the electric field ( $E_0$ ) drives an enhanced northward Hall current ( $J_H^0$ ) and if this current cannot be totally closed through the magnetosphere a southward polarization electric field ( $E_p$ ) is set up. This polarization field drives a secondary westward Hall current ( $J_H^p$ ) that adds to the westward component of the primary Pedersen current ( $J_p^0$ ), thus constituting a westward Cowling current (Figures 8-13 and 8-14).

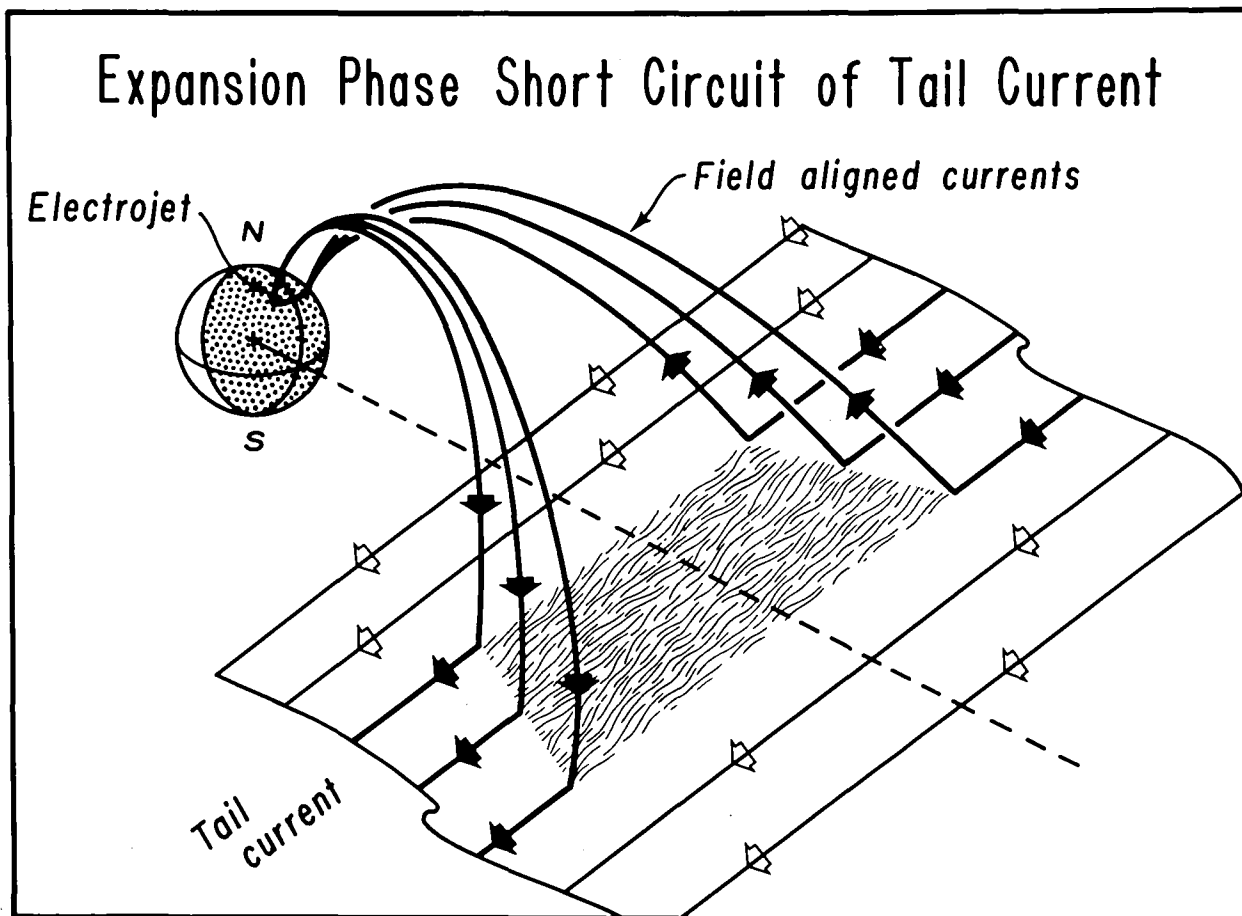


Figure 8-12. Schematic drawing of current flow on the nightside of the inner magnetosphere during a substorm. Current flowing across the tail has been diverted into the ionosphere along field lines to form the substorm electrojet. (From McPherron et al., 1973).

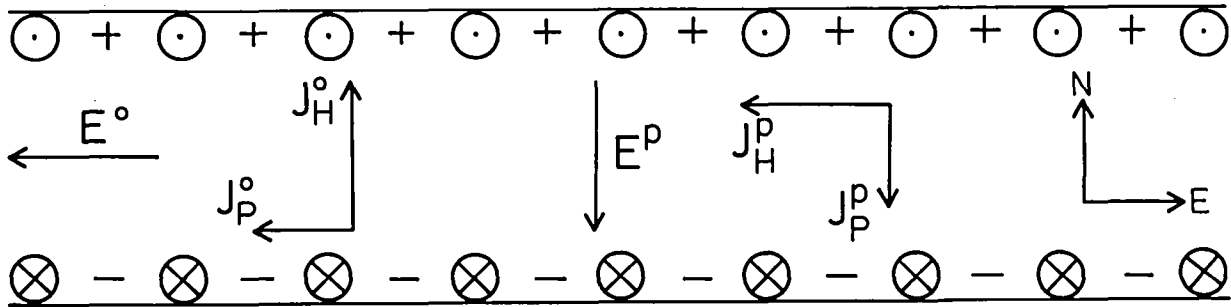


Figure 8-13. Generation of a polarization electric field and an incomplete Cowling channel in a region of enhanced ionization (after Bostrom, 1974).

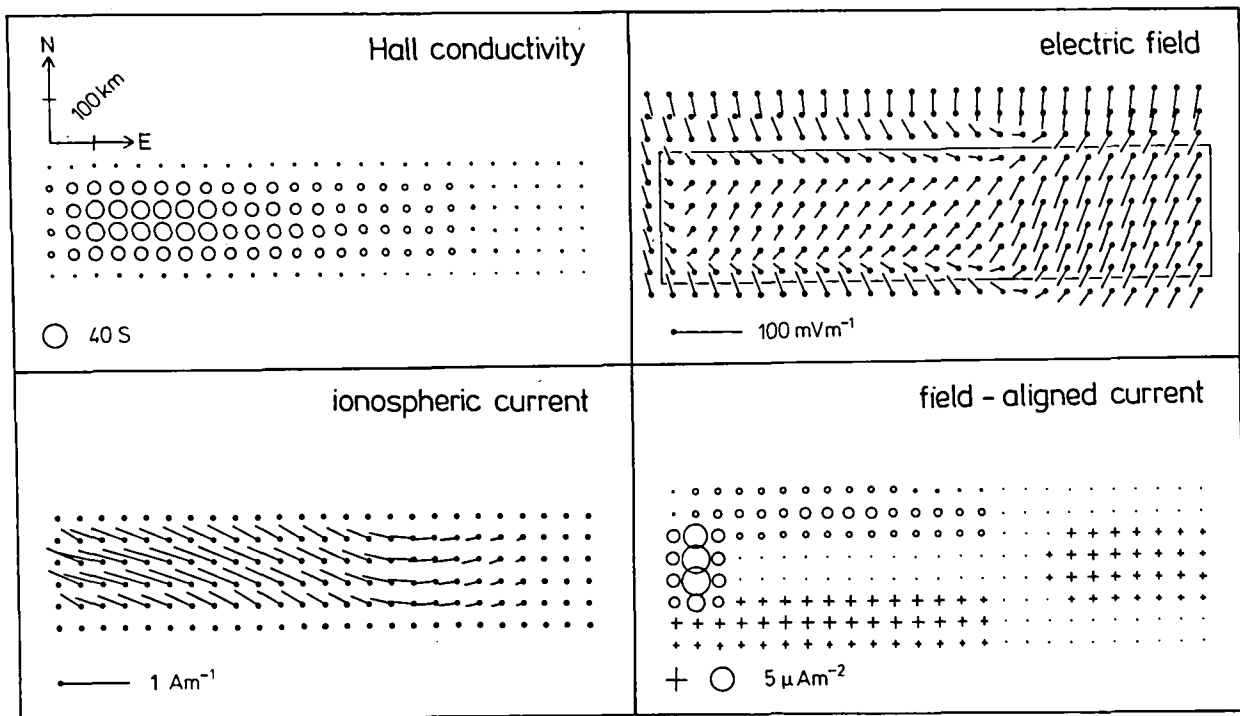


Figure 8-14. Spatial distribution of Hall conductivity, horizontal electric field vectors, and ionospheric and field-aligned current in the region of active break-up aurora around local magnetic midnight (after Baumjohann et al., 1981). The rectangle in the electric field panel frames the region of active aurora. Circles and crosses in the field-aligned current panel denote upward and downward direction of the field-aligned current flow.

Numerical model calculations based on observations during the initial brightening and peak development of the second, most prominent break-up, indicate that the northern and southern field-aligned currents are indeed very weak. Highly localized and intense upward field-aligned currents at the western edge of the active region and more wide spread and less intense downward currents in the east preserve the current continuity of the westward Cowling current and complete the near-Earth part of the substorm current wedge (Figure 8-14).

Inhester et al. (1981) have analyzed two-dimensional ground magnetic and ionospheric electric field observations during the passage of a westward traveling surge (WTS) in the late evening sector. On the basis of these measurements and optical auroral observations they modeled the three-dimensional current system in the vicinity of the WTS. Ahead of the surge, they found a south-eastward directed electric field. Additionally, they gave evidence for the existence of a south-westward directed electric field east of the surge. The latter may again be explained by the generation of polarization charges at the northern and southern boundaries of the higher conducting region east of the surge's head. Accordingly, the westward current in the higher conducting region east of the surge is again a Cowling current. Continuity of this current is conserved by a localized and intense upward field-aligned current in the head of the surge. No downward field-aligned current could be detected within the observation area (Figure 8-15).

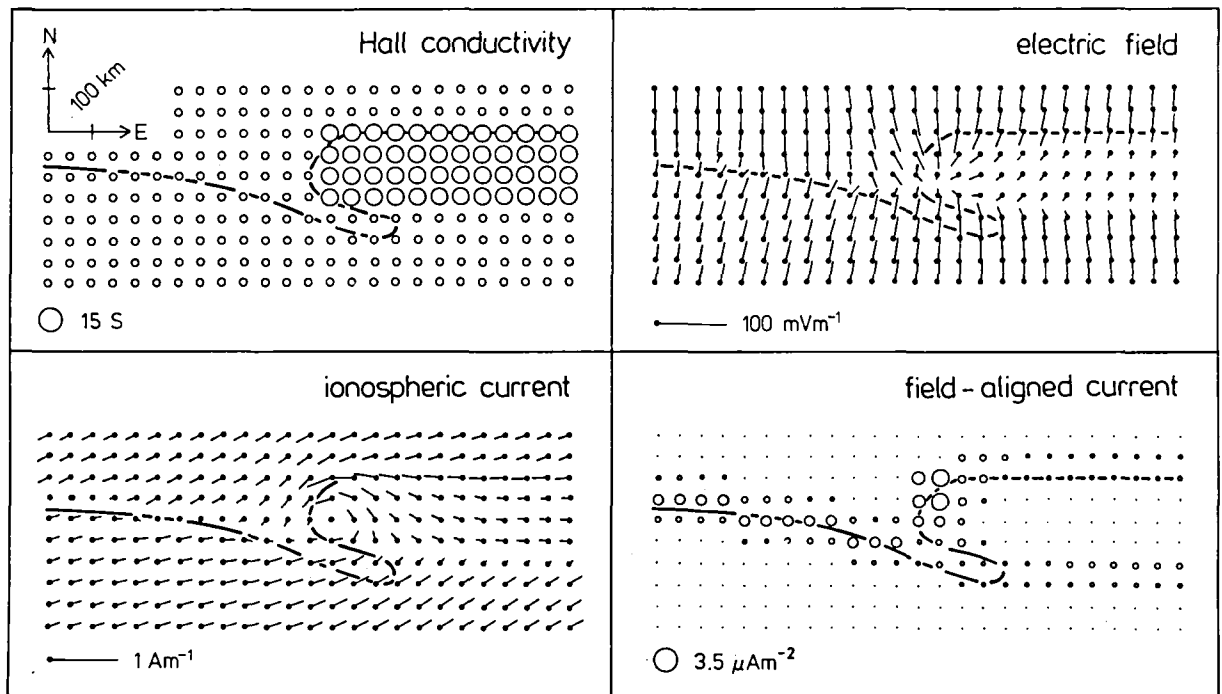


Figure 8-15. Same as Figure 8-14, but for a westward traveling surge in the early evening sector (after Inhester et al., 1981).



These two current models and also similar models proposed by Rostoker and Hughes (1979), Tighe and Rostoker (1981), Opgenoorth et al. (1983), and Marklund et al. (1983) show an important feature which may be relevant for the actual cross-tail current disruption process: the upward field-aligned current in the head of the westward traveling surge is highly confined and intense while the downward field-aligned current in the east is much more wide spread and less intense. Hence, the disruption of the tail current is not a symmetric process and it may even be possible that only the upward field-aligned current has its causal origin in the magnetospheric source region while the downward field-aligned current is just caused by the need to close the current loop.

## RECOVERY

The recovery is probably the least understood portion of the whole substorm process. The observations as stated in the Section on "Description of the Model" are very clear: the thickening of the plasma sheet, the increase in the  $B_z$  field component, and the poleward leap of the electrojet all accompany a decrease in the observed high-latitude currents and auroral luminosity. Low-altitude particle measurements (Winningham et al., 1975) also indicate that a narrow region of intense particle fluxes persist at the high-latitude border of the auroral region while the lower latitude portion of the newly formed thick plasma sheet decays. Our understanding of why these phenomena occur is far from clear.

### Possible Causes

The observations seem to indicate a model such as that reviewed in the "Description of the Model". The increased magnetotail  $B_z$  implies an increased number of closed field lines such as would result from reconnection. The enhanced volume of hot plasma would be the product of the reconnection process and the newly formed closed field lines provide the means for containing the plasma. But exactly why the plasma sheet remains thin until it abruptly recovers as a neutral line moves rapidly down the tail is not clear. Perhaps convection in the inner magnetosphere suddenly ceases in response to the imposed electric field controlled by the interplanetary magnetic field. If reconnected field lines continue to be produced by a relatively independent reconnection process in the tail, then these excess closed field lines would tend to "pile up" on the nightside. It would be surprising, however, if this chain of events could occur so rapidly that it would produce the abrupt thickening and rapid motion of the neutral line down the tail.

Swift (1982) has presented a quantitative model where thinning and subsequent thickening of the plasma sheet follow from the applied cross-tail electric field but without the necessity of a neutral line. This model requires no instability but it is not clear that it can reproduce the rapid tail changes that are observed.

### Quenching and Multiple Onsets

From the well known relationship between interplanetary  $B_z$  and geomagnetic activity it is clear that intervals of persistent northward field will tend to inhibit substorms. Indeed this relationship appears to hold true for individual case studies (e.g., Fairfield et al., 1981a); when the interplanetary  $B_z$  component turns and remains northward for an extended period geomagnetic activity decreases on a time scale on the order of an hour.

Although recovery is often a relatively well defined event in terms of the plasma sheet thickening and geomagnetic activity decrease, the same is not so of the energy density in the tail. Sometimes recovery seems to proceed until tail energy is exhausted and a "ground state" of the magnetosphere is reached. At other times, however, the activity decreases yet for some reason the tail

remains in an enhanced energy state. The reason for such occurrences is unknown, but is perhaps related to the conditions required for onsets. Thus again we see how interdependent are the various aspects of substorms -- the recovery of the substorm may be related to the processes responsible for onset.

## V. SUMMARY AND FUTURE DIRECTIONS

In this report we have provided a working definition of substorms, we have considered the primary relationships within the solar wind-magnetosphere-ionosphere system as they relate to substorms, and we have tried to clarify the role that we, as a committee, feel substorm studies play in the overall discipline of solar-terrestrial research. In pursuing the latter of these topics we have confronted the question of which aspects of magnetospheric activity are directly driven by the solar wind and which aspects of this activity represent an unloading process for previously stored solar wind energy.

In the Introduction we emphasized that substorm studies really comprise a global examination of the interrelationships between many detailed physical processes (see the reports of Working Groups 1, 2, 5, 7, and 11). Furthermore, we noted that substorms are a complex combination of the driven and unloading processes which have been described in some detail in the body of this report.

We feel that significant progress has been made in understanding substorm phenomenology and, to a large extent, in understanding the underlying theoretical (physical) mechanisms. Nonetheless, considerable disagreement and uncertainty remains. In studying substorms, we are struggling with a truly staggering global problem wherein experimental probing and/or theoretical modeling must deal with extremely large scale sizes. To a very large degree, the era in which a single scientific instrument on a single spacecraft can add substantially to the knowledge of substorm processes is past. We believe that a complete complement of instruments, from multiple spacecraft and ground-based observations if possible, should be employed to advance our understanding of magnetospheric dynamics. Such global measurements, combined with realistic global 3-D numerical simulations, probably hold the greatest promise of advancing our understanding significantly.

Substorm research has established that a variety of phenomena (e.g., auroral break-up, injections at synchronous orbit, and reconnection signatures in the distant magnetotail) occur at widely separated locations in association with the expansion phase onset. Further research -- both theoretical and observational -- is needed to gain a more detailed, quantitative understanding of the interrelationships among these various phenomena. In particular, the following questions need to be addressed.

- What is the nature of the energy transfer mechanisms between the presumed energy storage reservoir in the magnetotail and the various sites of energy dissipation? Proposed transfer mechanisms such as compression waves and field-aligned currents have been developed only qualitatively to date.
- How do these energy transfer mechanisms relate to and/or arise out of the reconnection process?

- What are the propagation time delays between the magnetotail and the inner magnetosphere/ionosphere, and what ramifications do these delays have for substorm dynamics?

A complete understanding of any given magnetospheric process requires that the role of that process in the global dynamics of the magnetosphere be identified. An essential aid to attaining this global perspective is the world-wide array of all-sky camera, magnetometers, riometers, and other monitors of magnetospheric processes. Note that all points in the outer magnetosphere -- roughly, beyond synchronous orbit -- have their magnetic footprints in auroral and polar regions. A suitably spaced array of stations in these regions thus provides a vital means of continuously monitoring the state of the magnetosphere on a global scale. To fully exploit this monitoring capability, further progress is needed in identifying the ground signatures of specific magnetospheric processes and in improving ground-based indices of energy flow and dissipation.

The auroral electrojet indices (AE, AU, AL) were first introduced by Davis and Sugiura (1966). Since then they have been extensively used as measures of the auroral electrojet intensity and magnetospheric substorm activity. Their accuracy and limitations have been discussed to some extent by several authors (cf., Rostoker, 1972; Allen and Kroehl, 1975), but there had been no systematic way to evaluate or calibrate them. The AE index is an important and useful index when it is used properly, but one must be carefully aware of, and take into account, its limitations.

During recent years, the AE(12) index has been used frequently for purposes which require accuracy well above that of the presently available index. In particular, it has been a general practice to assume that a poor correlation in the correlation studies between solar wind parameters and AE(12) is due to a poor choice of the particular solar wind parameter combination, rather than to the inaccuracy of AE(12). Since the occurrence frequency of the AE and AL indices peaks at  $\sim 100$  nT and since AE(12)  $< \sim 250$  nT may not be very reliable, one must be aware that poor correlations may not be entirely due to a poor choice of particular solar wind parameters. We firmly believe that future correlation studies between the AE index and solar wind quantities must be based on an improved AE index. Moreover it is crucial to continually improve geomagnetic indices and estimates of the global energy dissipation rates for there to be new advances in this particular area of magnetospheric substorm studies.

A specific problem (which we identify with present indices) is that there are indications that the current systems associated with the driven and unloading substorm processes have different morphologies and different spatial locations (Baumjohann and Kamide, 1983). The "unloading" current system apparently corresponds to the DP-1 current wedge system which creates the westward electrojet, while the "directly driven" system corresponds to a DP-2 system covering a larger area of the polar cap. If this is indeed the case then it would seem to be very important to distinguish between these current systems in future studies. Unfortunately, the AE index as presently formulated blurs this distinction since it normally will represent the weaker, more pervasive DP-2 system until it is preempted by the large effects of the westward electrojet. Thus, we again emphasize the need for greatly improved indicial formulation techniques in order to sort out the complexities of simultaneous driven and unloading substorm effects.

One of the most important problems in substorms studies has been to examine how the energy accumulated in the magnetotail varies during magnetospheric substorms. Satellite measurements of the magnetic field magnitude,  $B$ , provide a measure of the magnetic energy density ( $B^2/8\pi$ ) which may be characteristic of the entire volume of the magnetotail. However,  $B$  is sensitive to changes of the solar wind pressure and of the local plasma pressure. Furthermore such data usually only represent a single point measurement and the data are only available for limited periods. Therefore, it is important to look for a measure of tail magnetic energy, which is continuous and which is not subject to great uncertainty. One such new measure may be the

"diameter" of the auroral oval, since the total open flux in the tail is expected to be roughly proportional to  $B_G(d/2)^2$  (where  $B_G$  denotes the vertical component of the magnetic field in the polar cap region) and since an increase of the diameter  $d$  indicates an increase of the open flux.

In a related vein, a phenomenon that is observed with striking regularity in the magnetotail beyond  $r \approx 15 R_E$  is the thickening or recovery of the plasma sheet during the subsidence of negative bays at nightside auroral zone stations. In fact, the plasma sheet thickening sometimes seems to occur in very close coincidence with peak bay activity and with the beginning of final recovery of auroral zone bays and with the onset of bays at low polar cap locations. Thus it has been inferred that a "poleward leap" of the electrojet occurs in coincidence with the plasma sheet's recovery is ascribed to a sudden tailward retreat of the substorm neutral line to large distance. Lobe plasma, carried by convection into the neighborhood of the neutral line and jetting earthward from it, threaded with contracting, newly reconnected field lines is thought to be the material reconstituting the plasma sheet. Theoretical and modeling studies should be directed toward understanding the retreating neutral line phenomenon.

The conventional description of the auroral substorm contains no reference to the poleward leap, but has the auroras moving to high latitudes during the expansion phase and then retreating gradually to lower latitudes during a "recovery" phase. The later phase of substorms needs to be studied with the advanced equipment now available (magnetometer chains, auroral imaging from satellites, high-latitude auroral radars) to determine more precisely the features that are attendant upon the retreat of the neutral line and plasma sheet recovery.

The interface between the plasma sheet and the tail lobe is a region where very dynamic processes are often found including plasma jetting, beams of energetic particles, and intense plasma waves. These processes are particularly prominent as the interface (i.e., the boundary of the plasma sheet) surges over a satellite during plasma sheet recovery. It is not surprising, in the light of the neutral line model of substorms, that this interface should be such a site of activity since it contains magnetic field lines connected directly to an X-type neutral line. This interface should receive detailed study both to derive information it may reveal regarding the reconnection process and to further our understanding of the substorm sequence, particularly the phenomena of plasmoid ejection and the retreating substorm neutral line.

As we consider substorm energy dissipation, we realize that a reasonable first-order estimate can regularly be made for ring current injection rates, Joule heating rates, and auroral particle precipitation rates. Since these dissipation rates can be (at least roughly) estimated, they can be related to quantitative energy input estimates derived from dayside reconnection models, viscous interaction models, etc. A major remaining deficiency as we consider the partitioning of the available input energy is determining (on a regular and routine basis) how much energy simply escapes down the tail in the plasmoid structure. Methods should be developed to monitor this component of substorm energy dissipation.

In the area of injection and energization of hot plasma in the outer equatorial magnetosphere, several problems remain outstanding. Among these are:

- What is the role of global convection relative to the role of impulsive injections in the formation of the ring current?
- What is the nature of the earthward-propagating compressional waves observed in association with impulsive injection events? In particular, can present models of tail collapse give rise to such a wave?

- Are there other mechanisms that can be invoked, beside the observed compressional wave, to account for impulsive, dispersionless injection events?
- What is the relationship between the injected hot equatorial plasmas and the field-aligned currents which occur in association with substorm auroral displays?

To attack these and other questions in this area, we must experimentally determine the detailed spatial structures of the injection boundary/injection front. Furthermore, very detailed multiple-satellite timings of events in the tail, in the near-geostationary environment, and on the ground must be made in order to relate the earthward injection of hot plasma to substorm processes in the more distant tail. Detailed numerical models being developed to describe tail dynamical processes should be modified to include compressional wave propagation (if possible) and hot plasma injection characteristics. Finally, detailed statistical and event analyses need to be performed in order to relate hot plasma distributions to field-aligned current structures and ring current development.

Numerous other problems remain to be addressed. For example, the mechanisms for triggering plasma sheet instabilities that may be responsible for substorm expansion onsets remain largely undetermined. Detailed theoretical analyses of the possible role of ionospheric ions in lowering plasma sheet instability thresholds should be undertaken and appropriate experimental investigations should be performed. In fact, the entire area of heavy ions in magnetospheric dynamics is just opening up and, furthermore, the appreciation of positive feedback loops (as those involving the ionosphere) is only beginning to come about. As observations have improved and as theory has progressed (as discussed and demonstrated in this Workshop) there has developed a further realization of the very significant analogy which exists between the explosive acceleration of energetic particles in the Earth's magnetosphere during substorm onsets and on the Sun during solar flare onsets. As we compare and contrast various parts of the solar-terrestrial system, we should more fully exploit such analogous processes in sub-disciplines that are often far-removed from one another.

## VI. REFERENCES

- Aggson, T. L., J. P. Heppner, and N. C. Maynard, Observations of large magnetospheric electric fields during the onset phase of a substorm, J. Geophys. Res., **88**, 3981, 1983.
- Akasofu, S. -I., The development of the auroral substorm, Planet Space Sci., **12**, 273, 1964.
- Akasofu, S. -I., Physics of magnetospheric substorms, Dordrecht, D. Reidel, 1977.
- Akasofu, S. -I., Interplanetary energy flux associated with magnetospheric substorms, Planet. Space Sci., **27**, 425, 1979.
- Akasofu, S. -I., The solar wind-magnetosphere energy coupling and magnetospheric disturbances, Planet. Space Sci., **28**, 495, 1980.
- Akasofu, S. -I., Energy coupling between the solar wind and the magnetosphere, Space Sci. Rev., **28**, 121, 1981a.

- Akasofu, S. -I., Relationship between the AE and Dst indices during geomagnetic storms, J. Geophys. Res., **86**, 4820, 1981b.
- Akasofu, S. -I., Magnetospheric substorms: A newly emerging model, Planet. Space Sci., **29**, 1069, 1981c.
- Akasofu, S. -I., and J. R. Kan, Importance of initial ionospheric conductivity on substorm onset, Planet. Space Sci., **30**, 1315, 1982.
- Akasofu, S. -I. and C. -I. Meng, A study of polar magnetic substorms, J. Geophys. Res., **74**, 293, 1969.
- Akasofu, S. -I., Kimball, D. S., Meng, C. -I.: The dynamics of the aurora II: westward traveling surges. J. Atmos. Terr. Phys., **27**, 173-187, 1965.
- Akasofu, S. -I., A. T. Y. Lui, C. -I. Meng, and M. Haurwitz, Need for a three-dimensional analysis of magnetic fields in the magnetotail during substorms, Geophys. Res. Lett., **5**, 283, 1978.
- Akasofu, S. -I., Y. Kamide, J. R. Kan, L. C. Lee, and B. -H. Ahn, Power transmissions from the solar wind-magnetosphere dynamo to the magnetosphere and to the ionosphere: Analysis of the IMS Alaska meridian chain data, Planet. Space Sci., **29**, 721, 1981.
- Allen, J. H., and H. W. Kroehl, Spatial and temporal distributions of magnetic effects of auroral electrojets as derived from AE indices, J. Geophys. Res., **80**, 3667, 1975.
- Arnoldy, R. L., Signature in the interplanetary medium for substorms, J. Geophys. Res., **76**, 5189, 1971.
- Arnoldy, R. L., and K. W. Chan, Particle substorms observed at the geostationary orbit, J. Geophys. Res., **74**, 5019, 1969.
- Atkinson, G., The expansive phase of the magnetospheric substorm. In: Dynamics of magnetosphere, S. -I. Akasofu (ed.), pp. 461-481, Dordrecht, D. Reidel, 1979.
- Atkinson, G., Inverted V's and/or discrete arcs: A three-dimensional phenomenon at boundaries between magnetic flux tubes, J. Geophys. Res., **87**, 1528, 1982.
- Aubry, M. P., and R. L. McPherron, Magnetotail changes in relation to the solar wind magnetic field and magnetospheric substorms, J. Geophys. Res., **76**, 4381, 1971.
- Baker, D. N., and E. C. Stone, Energetic electron anisotropies in the magnetotail: identification of open and closed field lines, Geophys. Res. Lett., **3**, 557, 1976.
- Baker, D. N., P. R. Higbie, E. W. Hones, Jr., and R. D. Belian, High-resolution energetic particle measurements at 6.6  $R_E$ , 3, Low-energy electron anisotropies and short-term substorm predictions, J. Geophys. Res., **83**, 4863, 1978.

- Baker, D. N., R. D. Belian, P. R. Higbie, and E. W. Hones, Jr., High-energy magnetospheric protons and their dependence on geomagnetic and interplanetary conditions, J. Geophys. Res., **84**, 7183, 1979a.
- Baker, D. N., P. R. Higbie, E. W. Hones, Jr., and R. D. Belian, The use of  $>30$  keV electron anisotropies at  $6.6 R_E$  to predict magnetospheric substorms, Proceedings of the International Solar-Terrestrial Predictions Workshop, Vol. IV, p. B-12, Boulder, CO, 1979b.
- Baker, D. N., E. W. Hones, Jr., J. B. Payne, and W. C. Feldman, A high-time resolution study of interplanetary parameter correlations with AE, Geophys. Res. Lett., **8**, 179, 1981a.
- Baker, D. N., E. W. Hones, Jr., P. R. Higbie, R. D. Belian, and P. Stauning, Global properties of the magnetosphere during a substorm growth phase: A case study, J. Geophys. Res., **86**, 8941, 1981b.
- Baker, D. N., T. A. Fritz, B. Wilken, P. R. Higbie, S. M. Kaye, M. G. Kivelson, T. E. Moore, W. Studemann, A. J. Masley, P. H. Smith, and A. L. Vampola, Observation and modeling of energetic particles at synchronous orbit on July 29, 1977, J. Geophys. Res., **87**, 5917, 1982a.
- Baker, D. N., E. W. Hones, Jr., D. T. Young, and J. Birn, The possible role of ionospheric oxygen in the initiation and development of plasma sheet instabilities, Geophys. Res. Letters, 1337, 1982b.
- Baker, D. N., R. D. Zwickl, S. J. Bame, E. W. Hones, Jr., B. T. Tsurutani, E. J. Smith, and S. -I. Akasofu, An ISEE-3 high time resolution study of interplanetary parameter correlations with magnetospheric activity, J. Geophys. Res., **88**, 6230, 1983.
- Barfield, J. N., D. E. DeForest, and D. J. Williams, Simultaneous observations of substorm electrons: Explorer 45 and ATS-5, J. Geophys. Res., **82**, 531, 1977.
- Bargatze, L. F., D. N. Baker, R. L. McPherron, and E. W. Hones, Jr., Magnetospheric response for many levels of geomagnetic activity, J. Geophys. Res., submitted, 1983.
- Baumjohann, W., Ionospheric and field-aligned current systems in the auroral zone - A concise review, Adv. Space Res., **2(10)**, 55, 1983.
- Baumjohann, W., and Y. Kamide, Hemispherical joule heating and the AE indices, J. Geophys. Res., **88**, submitted, 1983.
- Baumjohann, W., R. J. Pellinen, H. J. Opgenoorth, E. Nielsen, Joint two-dimensional observations of ground magnetic and ionospheric electric fields associated with auroral zone currents. Current systems associated with auroral zone currents. Current systems associated with local auroral break-ups, Planet. Space Sci., **29**, 431, 1981.
- Belian, R. D., D. N. Baker, P. R. Higbie, and E. W. Hones, Jr., High-resolution energetic particle measurements at  $6.6 R_E$ , 2, High-energy proton drift echoes, J. Geophys. Res., **83**, 4857, 1978.

- Belian, R. D., D. N. Baker, E. W. Hones, Jr., P. R. Higbie, S. J. Bame, and J. R. Asbridge, Timing of energetic proton enhancements relative to magnetospheric substorm activity and its implication for substorm theories, J. Geophys. Res., **86**, 1415, 1981.
- Bieber, J. W., and E. C. Stone, Streaming energetic electrons in earth's magnetotail: evidence for substorm-associated magnetic reconnections, Geophys. Res. Lett., **7**, 945, 1980.
- Bieber, J. W., E. C. Stone, E. W. Hones, Jr., D. N. Baker, and S. J. Bame, Plasma behavior during energetic electron streaming events: further evidence for substorm-associated magnetic reconnection, Geophys. Res. Lett., **9**, 664, 1982.
- Birn, J., and E. W. Hones, Jr., Three-dimensional computer modeling of dynamic reconnection in the geomagnetic tail, J. Geophys. Res., **86**, 6802, 1981.
- Birn, J., and K. Schindler, Self-consistent theory of three-dimensional convection in the geomagnetic tail, J. Geophys. Res., in press, 1983.
- Bogott, F. H., and F. S. Mozer, Nightside energetic particle decreases at synchronous orbit, J. Geophys. Res., **79**, 8119, 1973.
- Bostrom, R., A model of the auroral electrojets, J. Geophys. Res., **69**, 4983, 1964.
- Bostrom, R., Ionosphere-magnetosphere coupling, in: Magnetospheric Physics, B. M. McCormac (ed.), pp. 45-59, Dordrecht, D. Reidel, 1974.
- Burlaga, L. F., and R. P. Lepping, The causes of recurrent geomagnetic storms, Planet. Space Sci., **25**, 1151, 1977.
- Burton, R. K., R. L. McPherron, and C. T. Russell, An empirical relationship between interplanetary condition and Dst, J. Geophys. Res., **80**, 4204, 1975.
- Caan, M. N., R. L. McPherron, and C. T. Russell, The statistical magnetic signature of magnetospheric substorms, Planet. Space Sci., **26**, 269, 1978.
- Caan, M. N., D. H. Fairfield, and E. W. Hones, Jr., Magnetic fields in flowing magnetotail plasmas and their significance for magnetic reconnection, J. Geophys. Res., **84**, 1971, 1979.
- Chao, J. K., J. R. Kan, A. T. Y. Lui, and S. -I. Akasofu, A model for thinning of the plasma sheet, Planet. Space Sci., **25**, 703, 1977.
- Chen, A. J., Penetration of low-energy protons deep into the magnetosphere, J. Geophys. Res., **75**, 2458, 1970.
- Chiu, Y. T., and J. M. Cornwall, Electrostatic model of a quiet auroral arc, J. Geophys. Res., **85**, 543, 1980.
- Clauer, C. R., and R. L. McPherron, Mapping the local time - universal time development of magnetospheric substorms using mid-latitude magnetic observations, J. Geophys. Res., **79**, 2811, 1974.



- Clauer, C. R., R. L. McPherron, C. Searls, and M. G. Kivelson, Solar wind control of auroral zone geomagnetic activity, Geophys. Res. Lett., **8**, 915, 1981.
- Clauer, C. R., R. L. McPherron, and C. Searls, Solar wind control of the low-latitude asymmetric magnetic disturbance field, J. Geophys. Res., **88**, 2123, 1983.
- Coppi, B., G. Laval, and R. Pellat, Dynamics of the geomagnetic tail, Phys. Rev. Lett., **16**, 1207, 1966.
- Coroniti, F. V., and C. F. Kennel, Polarization of the auroral electrojet, J. Geophys. Res., **77**, 2835, 1972.
- Crooker, N. U., J. Feynman, and J. T. Gosling, On the high correlation between long-term averages of solar wind speed and geomagnetic activity, J. Geophys. Res., **83**, 1933, 1977.
- Davis, T. N., and M. Sugiura, Auroral electrojet activity index AE and its universal time variations, J. Geophys. Res., **71**, 785, 1966.
- DeForest, S. E., and C. E. McIlwain, Plasma clouds in the magnetosphere, J. Geophys. Res., **76**, 3587, 1971.
- Dungey, J. R., Interplanetary magnetic field and the auroral zones, Phys. Rev. Lett., **6**, 47, 1961.
- Fairfield, D. H., and L. J. Cahill, Jr., Transition region magnetic field and polar magnetic disturbances, J. Geophys. Res., **71**, 155, 1966.
- Fairfield, D. H., R. P. Lepping, E. W. Hones, Jr., S. J. Bame, and J. R. Asbridge, Simultaneous measurements of magnetotail dynamics by IMP spacecraft, J. Geophys. Res., **86**, 1396, 1981a.
- Fairfield, D. H., E. W. Hones, Jr., and C. -I. Meng, Multiple crossing of a very thin plasma sheet in the earth's magnetotail, J. Geophys. Res., **86**, 11,189, 1981b.
- Foster, J. C., D. H. Fairfield, K. W. Ogilvie, and T. J. Rosenberg, Relationship of interplanetary parameters and occurrence of magnetospheric substorms, J. Geophys. Res., **76**, 6967, 1971.
- Fridman, M., and J. Lemaire, Relationship between auroral electron fluxes and field-aligned electric potential difference, J. Geophys. Res., **85**, 664, 1980.
- Galeev, A. A., and L. M. Zeleni, Tearing instability of plasma configurations, Sov. Phys. JETP, **43**, 1113, 1976.
- Galeev, A. A., F. V. Coroniti, and M. Ashour-Abdalla, Explosive tearing mode reconnection in the magnetosphere tail, Geophys. Res. Lett., **5**, 707, 1978.
- Goldstein, H., Die Ionen Tearing Instabilitat zwei-dimensionaler Schichtgleichgewichte und ihre Bedeutung fur die Dynamik der Magnetosphere, Ph.D. thesis, Ruhr-Universitat Bochum, Fed. Rep. Germany, 1981.

- Harel, M., R. A. Wolf, R. W. Spiro, P. H. Reiff, C. -K. Chen, W. J. Burke, F. J. Rich, and M. Smiddy, Quantitative simulation of a magnetospheric substorm, 2. Comparison with observations, J. Geophys. Res., **86**, 2242, 1981.
- Heppner, J. P., The Harang discontinuity in auroral belt ionospheric currents, Geophys. Publ., **29**, 105, 1972.
- Holzer, R. E., and J. A. Slavin, Magnetic flux transfer associated with expansions and contractions of the dayside magnetosphere, J. Geophys. Res., **83**, 3831, 1978.
- Holzer, R. E., and J. A. Slavin, An evaluation of three predictors of geomagnetic activity, J. Geophys. Res., **87**, 2558, 1982.
- Hones, E. W., Jr., The magnetotail: its generation and dissipation, in Physics of Solar Planetary Environments (D. J. Williams, ed.) p. 558, AGU Press, 1976.
- Hones, E. W., Jr., Substorm processes in the magnetotail: comments on "On hot tenuous plasma, fireballs, and boundary layers in the earth's magnetotail," by L. A. Frank, K. L. Ackerson, and R. P. Lepping, J. Geophys. Res., **82**, 5633, 1977.
- Hones, E. W., Jr., Transient phenomena in the magnetotail and their relation to substorms, Space Sci. Rev., **23**, 393, 1979.
- Hones, E. W., Jr., Plasma flow in the magnetotail and its implications for substorm theories, in Dynamics of the Magnetosphere, (S. -I. Akasofu, ed.) p. 545, 1980.
- Hones, E. W., Jr., and K. Schindler, Magnetotail plasma flow during substorms: a survey with IMP 6 and IMP 8 satellites, J. Geophys. Res., **84**, 7155, 1979.
- Hones, E. W., Jr., J. R. Asbridge, S. J. Bame, and I. B. Strong, Outward flow of plasma in the magnetotail following geomagnetic bays, J. Geophys. Res., **72**, 5879, 1967.
- Hones, E. W., Jr., J. R. Asbridge, S. J. Bame, and S. Singer, Substorm variations of the magnetotail plasma sheet from  $X_{SM} \sim -6 R_E$  to  $X_{SM} \sim -60 R_E$ , J. Geophys. Res., **78**, 109, 1973.
- Hones, E. W., Jr., S. J. Bame, and J. R. Asbridge, Proton flow measurements in the magnetotail plasma sheet made with IMP 6, J. Geophys. Res., **81**, 227, 1976.
- Hones, E. W., Jr., J. Birn, S. J. Bame, G. Paschmann, and C. T. Russell, On the three-dimensional magnetic structure of the plasmoid created in the magnetotail at substorm onset, Geophys. Res. Lett., **9**, 203, 1982.
- Inhester, B. W. Baumjohann, R. A. Greenwald, and E. Nielsen, Joint two-dimensional observations of ground magnetic and ionospheric electric field associated with auroral zone currents. 3. Three-dimensional currents associated with a westward travelling surge, J. Geophys. Res., **49**, 155, 1981.

- Iyemori, T., H. Maeda, and T. Kamei, Impulse response of geomagnetic indices to interplanetary magnetic field, J. Geomag. Geoelectr., **31**, 1, 1979.
- Kan, J. R., Towards a unified theory of discrete auroras, Space Sci. Rev., **31**, 71, 1982.
- Kan, J. R., and L. C. Lee, Theory of imperfect magnetosphere-ionosphere coupling, Geophys. Res. Lett., **7**, 633, 1980.
- Kan, J. R., L. C. Lee, and S. -I. Akasofu, The energy coupling function and the power generated by the solar wind-magnetosphere dynamo, Planet. Space Sci., **28**, 823, 1980.
- Kaye, S. M., and M. G. Kivelson, Time dependent convection electric fields and plasma injections, J. Geophys. Res., **84**, 4183, 1979.
- Kennel, C. F., Consequences of a magnetospheric plasma, Rev. Geophys. Sp. Phys., **7**, 379, 1969.
- Kisabeth, J. L., and G. Rostoker, The expansive phase of polar magnetic substorms, 1, The development of the substorm perturbation sequence, J. Geophys. Res., **79**, 972, 1974.
- Kivelson, M. G., S. M. Kaye, and D. J. Southwood, The physics of plasma injection events, Dynamics of the Magnetosphere, ed., S. -I. Akasofu, D. Reidel Publishing Co., Dordrecht - Holland, p. 385, 1980.
- Knight, S., Parallel electric fields, Planet. Space Sci., **21**, 741, 1973.
- Kokubun, S., R. L. McPherron, and C. T. Russell, Triggering of substorms by solar wind discontinuities, J. Geophys. Res., **82**, 74, 1977.
- Konradi, A., C. L. Semar, and T. A. Fritz, Substorm-injected protons and electrons and the injection boundary model, J. Geophys. Res., **80**, 543, 1975.
- Lezniak, T. W., and J. R. Winckler, Experimental study of magnetospheric motions and the acceleration of energetic electrons during substorms, J. Geophys. Res., **75**, 7075, 1970.
- Lui, A. T. Y., and J. R. Burrows, On the location of auroral arcs near substorm onsets, J. Geophys. Res., **83**, 3342, 1978.
- Lui, A. T. Y., E. W. Hones, Jr., D. Venkatesan, S. -I. Akasofu, and S. J. Bame, Complete plasma dropouts at Vela satellites during thinning of the plasma sheet, J. Geophys. Res., **80**, 4649, 1975.
- Lui, A. T. Y., C. -I. Meng, and S. -I. Akasofu, Search for the magnetic neutral line in the near-earth plasma sheet 1. Critical reexamination of earlier studies on magnetic field observations, J. Geophys. Res., **81**, 5943, 1976.
- Lui, A. T. Y., C. -I. Meng, and S. -I. Akasofu, Search for the magnetic neutral line in the near-earth plasma sheet 2. Systematic study of Imp 6 magnetic field observations, J. Geophys. Res., **82**, 1547, 1977a.
- Lui, A. T. Y., L. A. Frank, K. L. Ackerson, C. -I. Meng, and S. -I. Akasofu, Systematic study of plasma flow during plasma sheet thinnings, J. Geophys. Res., **82**, 4815, 1977b.

- Lyons, L. R., Discrete aurora as the direct result of an inferred high-altitude generating potential distribution, J. Geophys. Res., **86**, 1, 1981.
- Lyons, L. R., and D. J. Williams, A source for the geomagnetic storm main phase ring current, J. Geophys. Res., **85**, 523, 1980.
- Marklund, G., W. Baumjohann, and I. Sandahl, Rocket and ground-based study of an auroral breakup event, Planet. Space Sci., **31**, 207, 1983.
- Mauk, B. H., and C. E. Mcllwain, Correlation of Kp with the substorm-injected plasma boundary, J. Geophys. Res., **79**, 3193, 1974.
- Mauk, B. H., and C. -I. Meng, Characterization of geostationary particle signatures based on the "injection boundary" model, J. Geophys. Res., **88**, 3055, 1983a.
- Mauk, B. H., and C. -I. Meng, Dynamical injections as the source of near geostationary, quiet time particle spatial boundaries, submitted to J. Geophys. Res., April 1983b.
- Mcllwain, C. E., Plasma convection in the vicinity of the geosynchronous orbit, Earth's Magnetospheric Processes, ed. by B. M. McCormac, p. 268, D. Reidel, Dordrecht, Netherlands, 1972.
- Mcllwain, C. E., Substorm injection boundaries, in Magnetospheric Physics, ed. by B. M. McCormac, D. Reidel Pub. Co., Dordrecht-Holland, p. 143, 1974.
- McPherron, R. L., Growth phase of magnetospheric substorms, J. Geophys. Res., **28**, 5592, 1970.
- McPherron, R. L., C. T. Russell, and M. P. Aubry, Satellite studies of magnetospheric substorms on August 15, 1968, 9. Phenomenological model for substorms, J. Geophys. Res., **78**, 3131, 1973.
- Meloni, A., A. Wolfe, and L. J. Lanzerotti, On the relationship between interplanetary quantities and the global auroral electrojet index, J. Geophys. Res., **87**, 119, 1982.
- Meng, C. -I., B. Tsurutani, K. Kawasaki, and S. -I. Akasofu, Cross-correlation analysis of the AE index and the interplanetary magnetic field  $B_z$  component, J. Geophys. Res., **78**, 617, 1973.
- Meng, C. -I., B. H. Mauk, and C. E. Mcllwain, Electron precipitation of evening diffuse aurora and its conjugate electron fluxes near the magnetic equator, J. Geophys. Res., **84**, 2545, 1979.
- Moore, T. E., R. L. Arnoldy, J. Feynman, and D. A. Hardy, Propagating substorm injection fronts, J. Geophys. Res., **86**, 6713, 1981.
- Murayama, T., T. Aoki, H. Nakai, and K. Hakamada, Empirical formula to relate the auroral electrojet intensity with interplanetary parameters, Planet. Space Sci., **28**, 803, 1980.

- Nagai, T., Local-time dependence of electron flux changes during substorms derived from multi-satellite observation at synchronous orbit, J. Geophys. Res., **87**, 3456, 1982.
- Nagai, T., D. N. Baker, and P. R. Higbie, Development of substorm activity in multiple onset substorms at synchronous orbit, J. Geophys. Res., in press, 1983.
- Nishida, A., and E. W. Hones, Jr., Association of plasma sheet thinning with neutral line formation in the magnetotail, J. Geophys. Res., **79**, 535, 1974.
- Nishida, A., and N. Nagayama, Synoptic survey for the neutral line in the magnetotail during the substorm expansion phase, J. Geophys. Res., **78**, 3782, 1973.
- Oguti, T., Electric coupling between the magnetosphere and the ionosphere as a cause of polar magnetic disturbances and auroral break-up, Cosmic Electrodyn., **2**, 164, 1971.
- Opgenoorth, H. J., R. J. Pellinen, W. Baumjohann, E. Nielsen, G. Marklund, L. Eliasson, Three-dimensional current flow and particle precipitation in a westward travelling surge (Observed during the Barium GEOS rocket experiment), J. Geophys. Res., **88**, 3138, 1983.
- Pedersen, A., R. Grard, K. Knott, D. Jones, and A. Gonfalone, Measurements of quasi-static electric fields between 3 and 7 earth radii on Geos-1, Space Sci. Rev., **22**, 333, 1978.
- Pedersen, A., R. Grard, K. Knott, D. Jones, A. Gonfalone, and U. Fahleson, Measurements of quasi-static electric fields between 3 and 7 earth radii on GEOS-1, in Advances in Magnetospheric Physics with GEOS-1 and ISEE, Ed. K. Knott, A. Durvey and K. Ogilvie, D. Reidel Publ. Co., Dordrecht-Holland, p. 13, 1979.
- Pellinen, R. J., and W. J. Heikkila, Energization of charged particles to high energies by induced substorm electric field within the magnetotail, J. Geophys. Res., **83**, 1544, 1978a.
- Pellinen, R. J., and W. J. Heikkila, Observations of auroral fading before breakup, J. Geophys. Res., **83**, 4207, 1978b.
- Pellinen, R. J., W. Baumjohann, W. J. Heikkila, V. A. Sergeev, A. G. Yahnin, G. Marklund, and A. O. Melnikov, Event study of substorm growth and trigger phases and their relation to the energy coupling between solar wind and magnetosphere, Planet Space Sci., **30**, 371, 1981.
- Pellinen, R. J., W. Baumjohann, W. J. Heikkila, V. A. Sergeev, A. G. Yahnin, G. Marklund, A. O. Melnikov: Event study on pre-substorm phases and their relation to the energy coupling between solar wind and magnetosphere, Planet. Space Sci., **30**, 371, 1982.
- Perreault, P., and S. -I. Akasofu, A study of geomagnetic storms, Geophys. J. R. Astr. Soc., **54**, 547, 1978.
- Pytte, T., R. L. McPherron, and S. Kokubun, The ground signatures of the expansion phase during multiple onset substorms, Planet. Space Sci., **24**, 1115, 1976a.
- Pytte, T., R. L. McPherron, M. G. Kivelson, H. I. West, Jr., and E. W. Hones, Jr., Multiple-satellite studies of magnetospheric substorms: Radial dynamics of the plasma sheet, J. Geophys. Res., **81**, 5921, 1976b.

- Pytte, T., R. L. McPherron, E. W. Hones, Jr., and H. I. West, Jr., Multiple-satellite studies of magnetospheric substorms: Distinction between solar magnetic substorm and convection driven negative bay, J. Geophys. Res., **83**, 663, 1978a.
- Pytte, T., R. L. McPherron, M. G. Kivelson, H. I. West, Jr., and E. W. Hones, Jr., Multiple-satellite studies of magnetospheric substorms: Plasma sheet recovery and the poleward leap of auroral zone activity, J. Geophys. Res., **83**, 5236, 1978b.
- Reiff, P. H., R. W. Spiro, and T. W. Hill, Dependence of polar cap potential drop on interplanetary parameters, J. Geophys. Res., **86**, 7639, 1981.
- Rostoker, G., Geomagnetic indices, Rev. Geophys. Space Phys., **10**, 935, 1972.
- Rostoker, G., and F. P. Camidge, Localized character of magnetotail fluctuations during polar magnetic substorms, J. Geophys. Res., **76**, 6944, 1971.
- Rostoker, G., and T. J. Hughes, A comprehensive model current system for high latitude magnetic activity - II. The substorm component, Geophys. J. R. Astr. Soc., **58**, 571, 1979.
- Rostoker, G., H. -L. Lam, and W. D. Hume, Response time of the magnetosphere to the interplanetary electric field, Canadian J. Phys., **50**, 544, 1972.
- Rostoker, G., S. -I. Akasofu, J. Foster, R. A. Greenwald, Y. Kamide, K. Kawasaki, A. T. Y. Lui, R. L. McPherron, and C. T. Russell, Magnetospheric substorm-Definition and signatures, J. Geophys. Res., **85**, 1663, 1980.
- Russell, C. T., and R. L. McPherron, The magnetotail and substorms, Space Sci. Rev., **15**, 205, 1973.
- Sarris, E. T., and W. I. Axford, Energetic protons near the plasma sheet boundary, Nature, **77**, 460, 1979.
- Sato, T., T. Hayashi, J. Walker, and M. Ashaour-Abdalla, Neutral sheet current interruption and field-aligned current generation by three dimensional driven reconnection, Geophys. Res. Lett., **10**, 221, 1983.
- Schindler, K., A theory of the substorm mechanism, J. Geophys. Res., **79**, 2803, 1974.
- Schindler, K., and J. Birn, Self-consistent theory of time-dependent convection in the earth's magnetotail, J. Geophys. Res., **87**, 2263, 1982.
- Shepard, G. G., R. Bostrom, H. Derblom, C. -E. Falthammar, R. Gendrin, K. Kaila, A. Korth, A. Pedersen, R. Pellinen, and G. Wrenn, Plasma and field signatures of poleward propagating auroral precipitation observed at the foot of the GEOS-2 field line, J. Geophys. Res., **85**, 4587, 1980.
- Siscoe, G. L., and W. D. Cummings, On the cause of geomagnetic bays, Planet. Space Sci., **17**, 1795, 1969.

- Snyder, C. W., M. Neugebauer, and U. R. Rao, The solar wind velocity and its correlation with cosmic ray variations and with solar and geomagnetic activity, J. Geophys. Res., **68**, 6361, 1963.
- Speiser, T. W., and K. Schindler, Magnetospheric substorm models: Comparison with neutral sheet magnetic field observations, Astrophys. and Space Science, **77**, 443, 1981.
- Stenzel, R. L., W. Cushman, and N. Wild, Magnetic field line reconnection experiments, 4. Resistivity, heating, and energy flow, J. Geophys. Res., **87**, 111, 1982.
- Swift, D. W., Numerical simulation of the interaction of the plasma sheet with the lobes of the earth's magnetotail, J. Geophys. Res., **87**, 2287, 1982.
- Terasawa, T., Numerical study of explosive tearing model instability in one-component plasmas, J. Geophys. Res., **86**, 9007, 1981.
- Tighe, W. G., and G. Rostoker, Characteristics of westward travelling surges during magnetospheric substorms, J. Geophys. Res., **50**, 51, 1981.
- Tsurutani, B. T., and C. -I. Meng, Interplanetary magnetic field variations and substorm activity, J. Geophys. Res., **77**, 2964, 1972.
- Wiens, R. G., and G. Rostoker, Characteristics of the development of the westward electrojet during the expansion phase of magnetospheric substorms, J. Geophys. Res., **80**, 2109, 1975.
- Winningham, J. D., F. Vasuhara, S. -I. Akasofu, and W. J. Heikkila, The latitudinal morphology of 10 eV to 10 keV electron fluxes during magnetically quiet and disturbed times in the 2100 - 0300 MLT sector, J. Geophys. Res., **80**, 3148, 1975.
- Vasyliunas, V. M., A survey of low-energy electrons in the evening section of the magnetosphere with Ogo 1 and Ogo 3, J. Geophys. Res., **73**, 2839, 1968.
- Vasyliunas, V. M., Comparative magnetospheres, in Proceedings of the Theory Institute in Solar-Terrestrial Physics, Boston College, August 1982.





## CHAPTER 9

# IMPACT OF FLARES ON THE SOLAR TERRESTRIAL ENVIRONMENT

### WORKING GROUP MEMBERS

Lennard A. Fisk, Chairman  
*University of New Hampshire*

Roger L. Arnoldy  
*University of New Hampshire*

L. J. Lanzerotti  
*Bell Laboratories*

Robert Lin  
*University of California, Berkeley*

E. Oran  
*U. S. Naval Research Laboratory*

J. B. Reagan  
*Lockheed Palo Alto Research Lab*

Michael Schulz  
*Aerospace Corporation*

Bruce T. Tsurutani  
*Jet Propulsion Laboratory*

## **CHAPTER 9**

### **IMPACT OF FLARES ON THE SOLAR TERRESTRIAL ENVIRONMENT**

I.	Introduction	9-3
II.	The Flare Process	9-3
III.	Impacts on the Heliosphere	9-11
IV.	Impacts on the Magnetosphere	9-17
V.	Impacts on the Ionosphere and Atmosphere	9-20
VI.	Impacts on Technology	9-27
VII.	Summary	9-32
VIII.	References	9-33

## I. INTRODUCTION

The solar flare is one of the most dramatic phenomena in nature. A large amount of energy, in the form of photons and plasma and energetic particles, is released suddenly from the solar atmosphere and impacts the entire solar-terrestrial environment, from the Sun, through the heliosphere to the Earth and beyond. Flares generate natural phenomena such as shock waves and geomagnetic storms; flares affect man's technology and his ability to survive in space. It is the purpose of this chapter to assess our current understanding of these impacts.

A study of the impacts of solar flares encompasses many of the phenomena that are being considered in detail in other chapters of this volume. Solar flares are examples of particle acceleration (Chapter 2); solar flares affect solar wind-magnetosphere coupling (Chapter 5) and magnetosphere-ionosphere coupling (Chapter 7), etc. We summarize here only the effects which are the direct and unique results of solar flares. The reader is referred to other chapters for more complete and detailed discussions of the physics of the various regions of the solar-terrestrial environment.

We begin by summarizing the current knowledge of the flare process and its particle and photon emissions (Section II). We consider next the impact of flares on the different regions of the solar-terrestrial environment, in order of their distance from the Sun to Earth: the heliosphere (Section III), the magnetosphere (Section IV), and the ionosphere and atmosphere (Section V). We describe finally (Section VI) the effects of flares on modern technology and manned spaceflight. In all cases, we have attempted to distinguish clearly among what is known, what is suspected, and what is unknown. In Section VII we summarize the important research areas where we feel concentrated research in the next few years should be pursued.

## II. THE FLARE PROCESS

In this section we review briefly the current understanding of the flare process and its photon and particle emission. For a more detailed discussion and a more complete reference list, the reader is referred to several excellent review articles that have appeared recently (Ramaty et al., 1979; Forman et al., 1983; Fan et al., 1983).

Flares release  $\sim 10^{29}$  to  $\sim 10^{32}$  erg of energy in a time scale of from  $\sim 10$ s to about thirty minutes. Small subflares may occur about once an hour near the maximum of the solar cycle. The frequency of occurrence of the largest flares ( $\sim 10^{31}$ - $10^{32}$  erg of energy), which produce strong solar terrestrial effects, is illustrated in Figure 9-1 for the last two solar cycles. Note that the very largest events have the tendency to occur in the descending phases of the solar activity cycle, rather than at maximum. The maximum rate of energy release in the largest flares is  $\sim 10^{30}$  erg/s.

In flares, an impulsive phase of  $\sim 10$  to a few hundred seconds is often observed first, followed by a gradual phase lasting 10-30 minutes. The impulsive phase tends to be dominated by non-thermal phenomena while the gradual phase is predominantly thermal in character.

Flares seen on the disk of the Sun have a typical scale size of  $\sim 10^3$ - $10^5$  km. The flare is observed to be most prominent in the chromosphere and corona, with photospheric emission (e.g., white-light flares) occurring rarely. Flares are located in regions of strong magnetic fields,  $\sim 10^2$ - $10^3$  Gauss, usually with complex, mixed magnetic polarity.

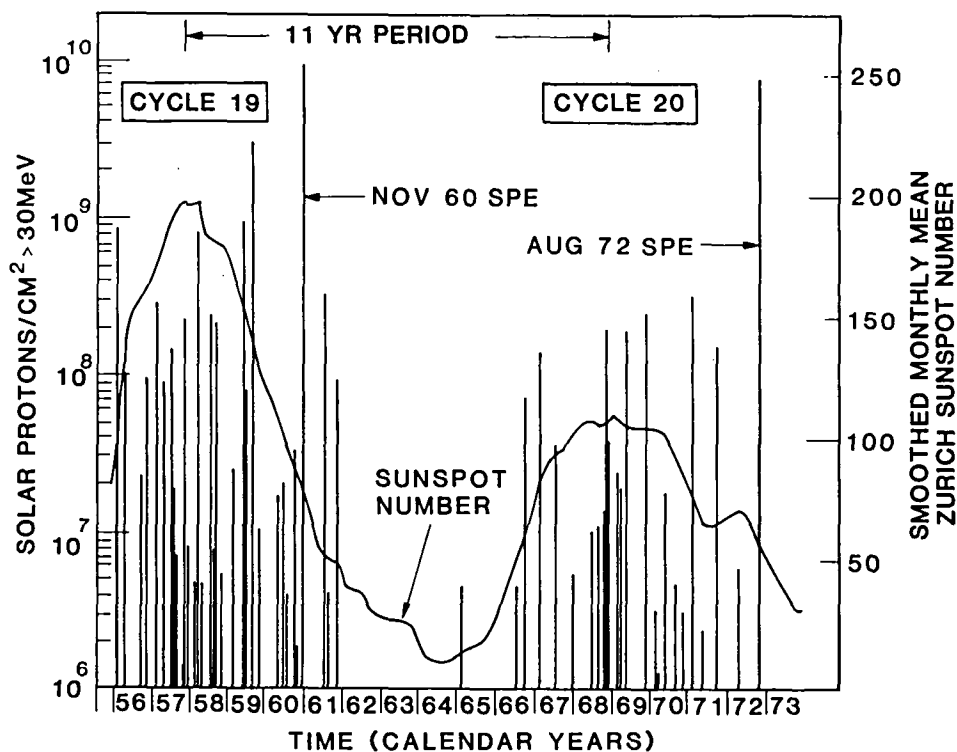


Figure 9-1. Occurrence frequency of solar proton events (>30 MeV) near Earth during solar cycles 19 and 20. The smoothed monthly-mean Zurich sunspot number is shown for reference (adapted from King, 1974 with permission from the J. Spacecr. Rockets)

## THE IMPULSIVE PHASE

Energetic particles are first accelerated in solar flares in the impulsive phase in a sudden release of magnetic energy stored in the solar atmosphere. Electrons with energies of 10-100 keV are the primary product of this acceleration; protons or other ions, with, e.g., similar velocities to the electrons are not accelerated in comparable numbers (Lin, 1974; Chupp 1983). The acceleration of comparable number of 10-100 keV protons, however, cannot be excluded; these energy nuclei do not have an easily observed radiation signature.

The accelerated electrons at the flare site yield a variety of forms of electromagnetic radiation: hard X-rays due to bremsstrahlung, impulsive microwave bursts due to gyrosynchrotron radiation, and type III bursts due to the conversion of Langmuir waves excited by electron beams into electromagnetic emission. The energy contained in the 10-100 keV electrons can be determined if it is assumed that the observed hard X-rays are produced by fast electrons in a cold medium. For many flares this energy can be comparable to the total energy released in the flare; i.e., in simple terms the principal occurrence in a solar flare is to take energy stored in the solar magnetic field and convert it into 10-100 keV electrons (Lin and Hudson, 1976).

An alternative to the nonthermal origin for the hard X-rays is that this emission is the result of a very hot,  $10^8$ - $10^9$ K thermal plasma. Then electron-electron Coulomb collisions will result in exchanges of energy only among the hot electrons without any net collisional energy loss. Provided the hot plasma could be confined, the energy requirements for the energetic electrons

might be drastically reduced (see discussion in Ramaty et al., 1979). However, recent Solar Maximum Mission (SMM) observations of hard X-ray images from flares appear to support the non-thermal interpretation (Hoyng et al., 1981, Duijveman et al., 1982). In particular, during the hard X-ray impulsive spikes of the flare of May 21, 1980 (Figure 9-2), the 10-30 keV hard X-rays came from simultaneously brightening footpoints (Figure 9-3); footpoints a and b in the first spike and a and c in the second. Upon comparison with the SMM UV spectrometer they find that these footpoints coincide with flare brightenings in Mg II and H lines (Figure 9-4) which are formed in a partially ionized medium. Thus the ambient temperature in the hard X-ray emitting region must be much lower than the fast electron energy, as required for nonthermal bremsstrahlung emission. Hoyng et al. (1981) also point out that thermal models for the hard X-ray emission where the very hot plasma is generated at the top of a flare loop and then energy is thermally conducted down to the footpoints can also be ruled out. Such conduction fronts would take  $\sim 10$ s to travel down, while no sources were observed located between the footpoints prior to the footpoint brightening themselves. It should be noted that only a few flares observed by SMM were sufficiently strong and spatially extended as observed in 16 keV X-rays for the above analysis to be performed.

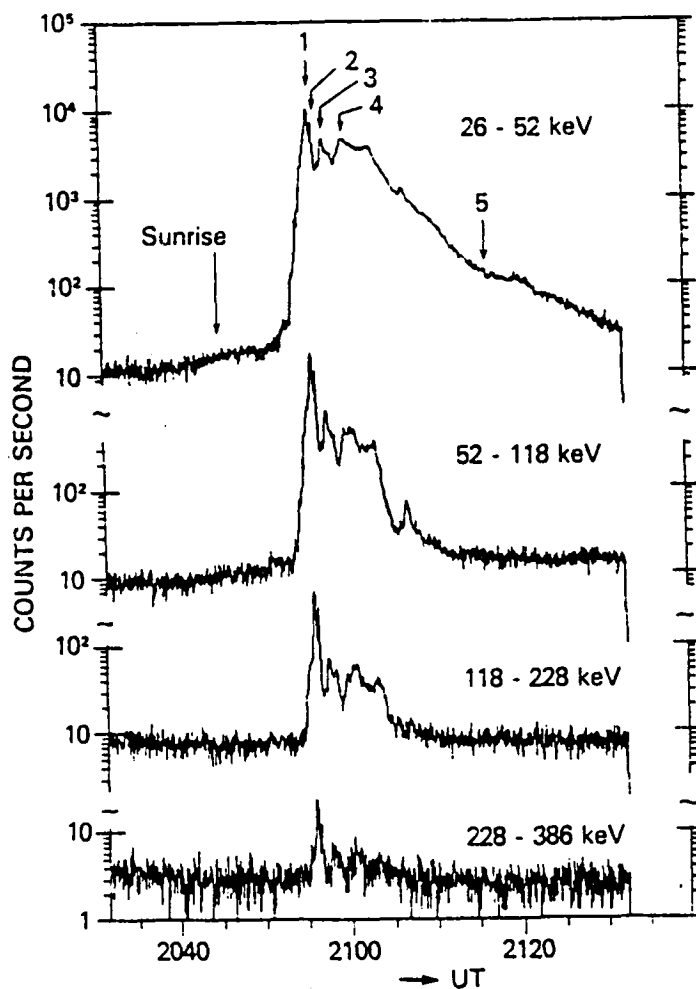


Figure 9-2. Hard X-ray time profiles as observed by the HXRBS instrument on SMM on 1980 May 21 (Hoyng et al., 1981.)

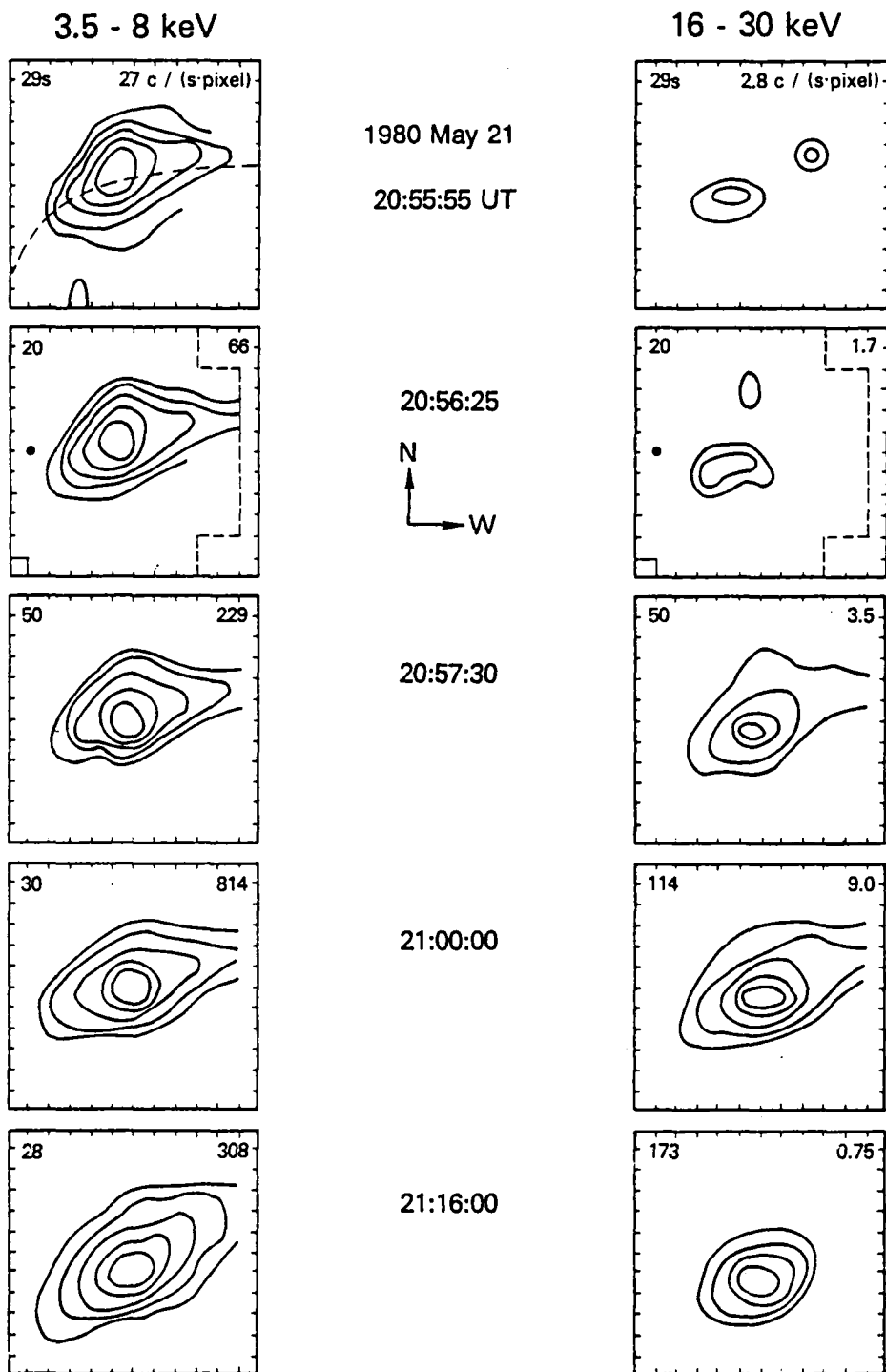


Figure 9-3. X-ray contour plots from the HXIS instrument on SMM at the times indicated by arrows in Figure 9-2. The contours are at fractions of  $3/4$ ,  $1/2$ ,  $1/4$ ,  $1/8$ , and  $1/16$  of the maximum brightness. Both contour plots at  $20^{\text{h}}56^{\text{m}}25^{\text{s}}$  show the center ( $\cdot$ ) and the boundary (---) of the HXIS fine field of view. The marks on the axes are  $8''$  apart (Hoyng et al., 1981.)

## FLARE KERNELS

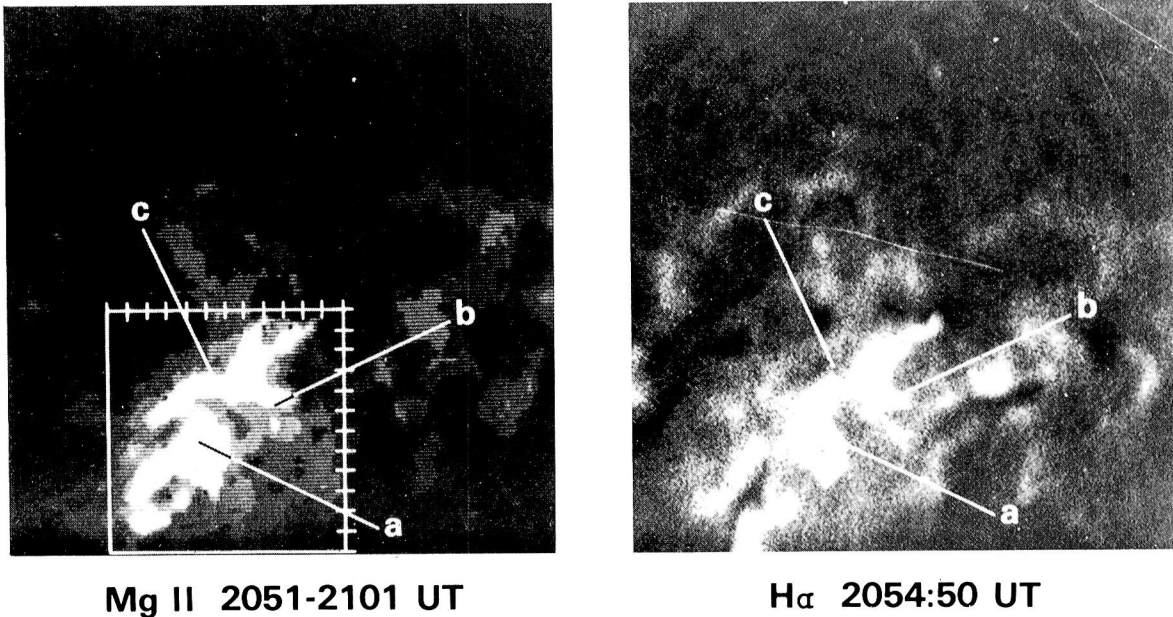


Figure 9-4. The flare kernels of 1980 May 21. Left: UVSP Mg II spectroheliogram  $\gamma 2795$  measuring about  $4' \times 4'$ . Pointers mark the centers of the three brightest  $8'' \times 8''$  HXIS pixels involved in the 16-30 keV emission around  $20^{\text{h}}55^{\text{m}}50^{\text{s}}$  and  $20^{\text{h}}56^{\text{m}}25^{\text{s}}$ . Right: flare patrol  $\text{H}\alpha$  frame at  $20^{\text{h}}54^{\text{m}}50^{\text{s}}$ . Points a, b, and c indicate the centers of the three most important HXIS pixels involved in the 16-30 keV emission around  $20^{\text{h}}56^{\text{m}}$  (Hoyng et al., 1981.)

Most other flare phenomena can be attributed to the interaction of the accelerated particles with the solar atmosphere. Optical, UV, EUV, and soft X-ray emission is due to heating of the chromosphere and corona to temperatures varying from  $\sim 10^4$  to  $3 \times 10^7 \text{ K}$ , in ambient densities of  $\sim 10^8$  to  $\sim 10^{12} \text{ cm}^{-3}$ . At lower levels in the solar atmosphere the energy input will be dissipated by radiation. At the higher levels the ambient gas will rapidly heat up to form the soft X-ray emitting  $\sim 10^7 \text{ K}$  flare plasma.

The efficiency and rapidity with which acceleration occurs during the impulsive stage suggests that it is the result of a direct electric field. The expectation is that the motion of magnetic fields in the solar atmosphere will lead to twisted field structures and sharp interfaces with accompanying high current densities. The sudden dissipation of these currents, by various plasma mechanisms such as the generation of anomalous resistivity or reconnection, will relieve the magnetic stresses and generate a large-scale electric field that can efficiently and rapidly accelerate thermal electrons to energies  $\sim 10$ -100 keV (e.g., D. Smith, 1979; Syrovatskii, 1981). It seems likely that these processes occur in the solar corona.

Approximately 0.1% to 1% of the electrons accelerated during the impulsive phase escape into the solar wind. The differential intensity spectrum is a power law in kinetic energy with spectral index typically between -2.5 and -4, up to energies  $\sim 100 \text{ keV}$ . If there is no further acceleration in the flare (see below), the spectrum decreases rapidly beyond  $\sim 100 \text{ keV}$  (Lin, 1974).

## SUBSEQUENT ACCELERATION

In many flares, particularly the larger ones, there is evidence for subsequent acceleration, during which electrons with energies  $>100$  keV and energetic ions are accelerated. The acceleration of relativistic electrons and ions is accompanied by type II radio bursts, which are expected to result from the conversion of Langmuir waves generated by low energy electrons accelerated near shock fronts (Wild et al., 1963). The relativistic electrons yield type IV microwave bursts and moving type IV bursts due to synchrotron radiation, and the gamma-ray continuum due to bremsstrahlung.

Approximately 0.1% to 1% of the electrons then escape into the solar wind where they are observed to have a double power law spectrum: at energies  $<100$  keV the spectral index of the differential intensity spectrum is typically  $-0.6$  to  $-2.0$ , which is systematically harder than in events with only an impulsive phase; at energies  $>100$  keV, the spectral index is  $-2.4$  to  $-4.3$  (Lin et al., 1982).

Energetic ions accelerated to energies  $>10$  MeV/nucleon collide with ambient gas to yield gamma-ray lines resulting from nuclear reactions. The time-profile of the gamma-ray lines is similar to that of the microwave emission and gamma-ray continuum produced by the relativistic electrons, and reaches maximum intensity at a few seconds to a few minutes after the impulsive phase. Energetic neutrons are also produced by the nuclear reactions generated by the accelerated ions and have been observed near Earth on SMM (Chupp, 1983). These observations imply that protons with energies  $\sim 1$  GeV are accelerated promptly in flares.

It is not clear whether the acceleration of relativistic electrons and energetic ions occurs in a stage physically distinct from the impulsive phase. In the large August 4, 1982 flare of the last solar cycle there appears to be a delay of  $\sim 1$  minute between the impulsive 10-100 keV hard X-ray burst and the bulk of the  $>300$  keV hard X-rays and nuclear gamma-ray production. However, many events observed by SMM during the present cycle (Chupp, 1983) show little ( $\leq$  a few seconds) or no delay between 10-100 keV hard X-rays and the gamma-ray line and continuum emission (Figure 9-5). Clearly, impulsive acceleration of electrons to MeV energies and ions to GeV energies can occur in time scales of seconds, but whether it occurs on this short time-scale in every flare will require higher sensitivity gamma-ray detectors than now available.

Energetic ions that escape into the solar wind have an elemental composition similar to that of the corona and solar wind (Cook et al., 1980), and charge states characteristic of that of average coronal temperatures (Sciambi et al., 1977). A range of temperatures for the source region from  $5 \cdot 10^5$  K to  $5 \cdot 10^6$  K appears to be necessary to account for the data with a typical temperature of  $<2 \cdot 10^6$  K required (Ma Sung et al., 1981a). A common feature in the accelerated ions is also the presence of  $\text{He}^+$  ( $\text{He}^+/\text{He}^{++} \sim 0.1$  at equal energy per nucleon), which must have formed in regions with temperature  $<10^5$  K. There is no measurable evidence at present of similar low charge states of heavier elements, e.g.,  $\text{O}^+$ , (Gloeckler et al., 1981).

The energetic ions that escape into the solar wind do not appear to have passed through extensive material during their acceleration. Studies of the effects of ionization place an upper limit to the path-length of  $\lesssim 3$  mg/cm<sup>2</sup> (Mason et al., 1983); studies of the effects of spallation place the limit at  $\lesssim 50$  mg/cm<sup>2</sup> (Cook et al., 1980).

Various analytic forms have been suggested to fit the spectra of the accelerated ions. McGuire et al. (1981) find that energetic H and He data observed in the solar wind can be fit by a differential intensity spectrum that is proportional to a Bessel function. Ramaty (1982) finds that an identical spectrum will also account for the energetic ions that yield the gamma-ray lines.



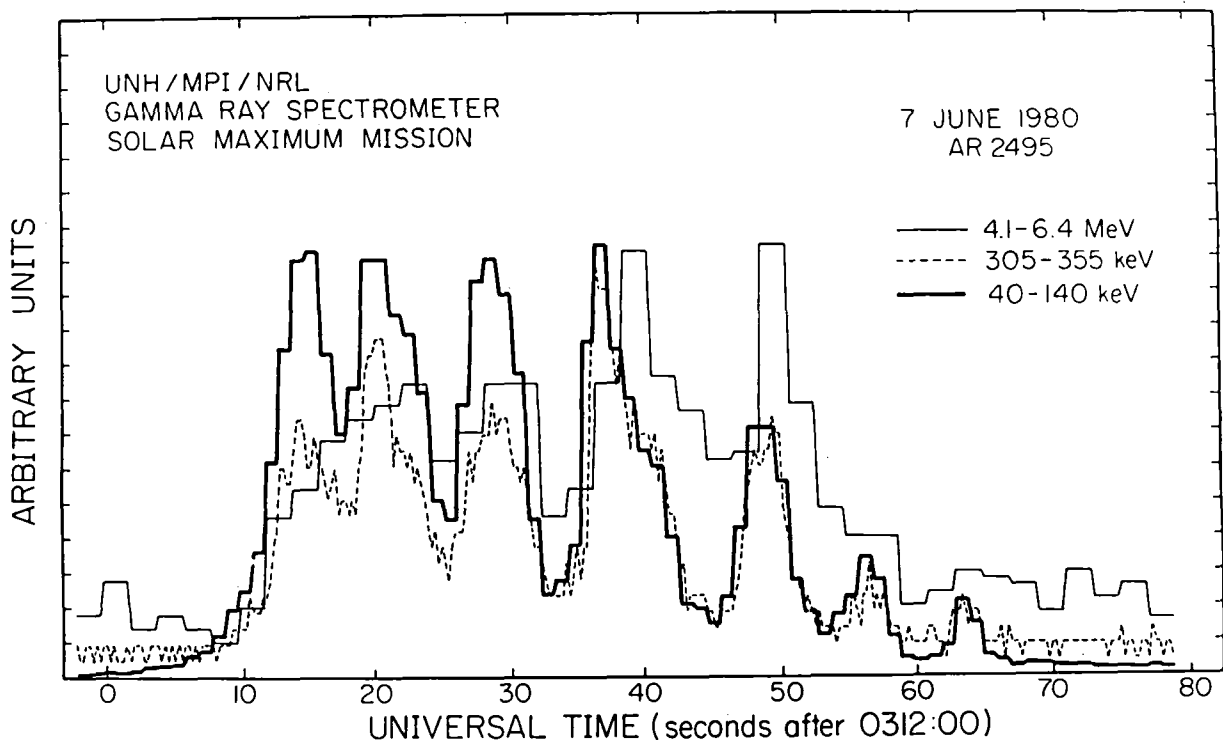


Figure 9-5. Counting rate versus time for various energy channels of the SMM Y-ray spectrometer during the Y-ray line flare at 0312 UT on 7 June 1980. (From Chapp, 1983)

Gloeckler et al. (1981) find that heavy ion data all have the identical spectral shape if the spectra are portrayed as a distribution function  $f \propto \exp(-vR^n/\eta_0)$  where  $v$  is particle speed,  $R$  is particle rigidity, and  $\eta_0$  and  $n$  are adjustable parameters. This latter form also accounts for the absence of low charge state heavy ions (e.g.,  $O^+$ ) when  $He^+$  is observed. The rigidities of the heavy ions are sufficiently high so that the fluxes predicted by this form for  $f$  are below current detectable thresholds.

In large flares substantial mass motions in the chromosphere and particularly in the corona are observed. The coronal mass ejections into the solar wind from a large flare often carry a large fraction of the total flare energy. The velocities for these flare-associated ejecta are higher than for non-flare associated coronal transients. The flare ejecta very likely form the driver gas for corona and solar wind shock waves.

The sudden release of energy into the plasma of the solar atmosphere by the flare and the subsequent mass motions may generate a shock wave in the corona which can be responsible for the acceleration of high energy electrons and energetic ions in the second stage of the acceleration process. The shock wave will yield, through the energetic electrons it generates, the observed type II radio bursts. The shock wave will pick up and accelerate ions from the surrounding corona, and thus yield accelerated ions with coronal composition and charge states. The acceleration must occur sufficiently far out in the corona and/or with access to magnetic fields that are open to the solar wind, so that the ions can escape without passing through appreciable matter. It is necessary also that the coronal material encountered by the shock wave is inhomogeneous in temperature, to account for the range of temperatures in the source region that are implied by charge-state measurements, and that the shock encounters cool coronal loops containing  $\text{He}^+$  (Fan et al., 1983).

The simplest form of a shock acceleration model is a first-order Fermi or compressive process in which particles are accelerated by scattering back and forth between the flows that are converging towards the shock front. As noted by Lee and Fisk (1982), such a mechanism in the presence of adiabatic cooling from a diverging flow upstream from the shock will yield the spectral form  $f \propto \exp(-vR^n/\eta_0)$  observed by Gloeckler et al. (1981). The shock wave may also generate turbulence which accelerates the particles stochastically. The Bessel function form of the spectrum used by McGuire et al. (1981) and Ramaty (1982) results from a highly simplified stochastic acceleration model.

Energetic ions observed in the solar wind occasionally exhibit compositional anomalies. The most dramatic example is  $^3\text{He}$ -rich flares where the  $^3\text{He}/^4\text{He}$  ratio at energies  $\sim 1$  MeV/nucleon can be enhanced by a factor  $\sim 10^3$  over expected solar abundances (e.g., Garrard et al., 1973). The heavy ions in  $^3\text{He}$ -rich flares can also be overabundant, e.g., Fe (Anglin et al., 1977) and appear to have unusual charge states, i.e.,  $\text{O}^{+5}$  and  $\text{Fe}^{+16-18}$  (Ma Sung et al., 1981b). Less dramatic compositional variations in other elements are also observed, particularly in small flares (Mason et al., 1980).

The compositional variations seen in small flares, particularly the dramatic events such as  $^3\text{He}$ -rich flares, are believed to result from plasma injection mechanisms, in which certain ion species in the thermal plasma are preferentially heated and thus preferentially injected into the flare acceleration process. The mechanism of Fisk (1978), which has recently been extended by Varvoglis and Papadopoulos (1983), uses an instability that excites electrostatic ion cyclotron waves to preferentially heat  $^3\text{He}^{+2}$ , and the accompanying heavy ions in  $^3\text{He}$ -rich flares, e.g.,  $\text{O}^{+5}$  and  $\text{Fe}^{+17}$ . Similar mechanisms may be able to account for all the compositional variation seen in small flares (Mason et al., 1980).

## OUTSTANDING QUESTIONS

There are numerous basic aspects of the flare process which remain unknown, despite extensive study in recent years. The exact location of the region where the energy is released is unclear, as are the conditions within this region, e.g., density, temperature, magnetic and electric fields, currents, etc. The mechanisms by which the energy is released and particles are accelerated are specified only in conceptual form. Flare geometries are intrinsically complex. The impulsive phase occurs because the magnetic field is forced into complicated geometries that must be

relieved. The shock wave generated by the impulsive phase propagates through and accelerates particles out of the inhomogeneous corona. Current theories address only basic concepts, which are demonstrated to occur in idealized situations. Theories are unable at present to model realistic coronal conditions and processes.

The physics by which a flare ejects material and forms a shock wave is not understood. A simple impulsive heating of the solar atmosphere probably does not account for this phenomenon. Flare regions contain strong magnetic fields and their realignment appears to be an essential aspect of the mass injection.

The relationship between the energetic ions that yield gamma-ray lines and those that escape into the solar wind is unclear. The former are produced shortly following the impulsive phase, and thus are near the heated flare site. The latter appear to originate in the quiet, surrounding corona. However, the spectra of the two particle populations appear to be similar. It may be that both populations are produced by the same shock mechanism but at different locations: the shock near the flare produces particles that are trapped in regions of relatively high densities and thus yield gamma rays; the shock in the outer corona produces particles that escape onto field lines leading into the solar wind. It may be also that the two populations are unrelated and produced by separate acceleration mechanisms.

No complete theory exists, in even idealized situations, which incorporates a preferential heating mechanism to explain compositional variations and flare acceleration mechanism. If the energetic ions seen in the small anomalous flares are produced by shock waves, then the preferential heating mechanism should be an intrinsic feature of the regions through which the shock propagates, or an intrinsic feature of the shock process itself.

One of the consequences of the inadequacy of our understanding of the flare process is that there is no reasonable expectation in the near future that a reliable predictive capability of flare occurrence can be developed. Such a capability will require a detailed understanding of the conditions in which flares occur, e.g., the magnetic field configuration, and an ability to observe when these conditions have developed to the stage where a flare can be triggered. It should be noted that when a flare occurs reasonable predictive capabilities from its radio emission are available to determine when large fluxes of energetic particle emissions are expected to strike Earth (Castelli and Guidice, 1972). However, these capabilities are currently available only for flares that occur on the visible disk of the Sun; spacecraft which make detailed observations of photon emissions from the invisible hemisphere of the Sun are not currently available.

### III. IMPACTS ON THE HELIOSPHERE

Of the many manifestations of solar flares, such as photon emission and the release of flare particles, the feature that propagates the slowest and has the longest delayed effects is the ejected plasma and fields. X-rays, UV and visible light propagate directly to the Earth, taking only eight minutes for the transit. The path length of the energetic protons, alphas and electrons is somewhat lengthened by their traversal along the spiraled Parker (1958) field lines, leading to delay times of tens of minutes to days, depending on the particle kinetic energy, and the solar longitude of the flare. In comparison, the ejected plasma and field propagate at relatively slow speeds

of 500-1000 km/sec and take two to three days to reach 1 AU. Although the effects of this portion of the flare are greatly delayed, they can frequently be substantial. Over fifty percent of the flare energy,  $10^{30}$ - $10^{32}$  erg, is associated with the ejected plasma and fields. During periods near solar maximum, there is an average of 3-4 of these solar flare shock events detected per month at the Earth, implying relatively frequent injection of energy into the solar-terrestrial environment.

The solar flare-associated shock and its ensuing high speed stream have four distinct effects on the solar-terrestrial environment: (1) dissipation of the high speed flow energy and consequential heating of the solar wind plasma; (2) energization of electrons to several tens of keV and ions to energies exceeding 1 MeV; (3) substantial decreases in ambient galactic cosmic ray flux caused by particle interaction with the shocks and the compressed downstream magnetic fields; and (4) geomagnetic activity, such as Sudden Impulses (SIs) and magnetic storms and substorms caused by the compressed plasma at the shock front and the distorted magnetic fields behind the shock, respectively. All four processes derive their energy from the directed kinetic energy of the high speed shock-associated solar wind plasma, decelerating the plasma in the process.

## SOLAR WIND HEATING

As a flare-associated shock propagates into the interplanetary medium, it heats and accelerates the upstream ambient plasma causing deceleration and dissipation of the directed high speed flow of the driver gas (Goldstein and Jokipii, 1977). At 1 AU, typical upstream temperatures are  $T_p = 8 \times 10^4$  K and  $T_e = 1.5 \times 10^5$  K whereas the downstream values behind typical shocks are  $2.6 \times 10^5$  K and  $2.2 \times 10^5$  K, respectively (Feldman et al., 1983). Long term effects are observed in the decrease in the velocity variances with increasing heliocentric distances (Collard et al., 1982; Smith and Barnes, 1983). Also, the proton temperature does not fall off adiabatically with distance, it has a  $r^{-2/3}$  dependence (Gazis and Lazarus, 1983). This lack of adiabaticity is believed to be due to the heating effects of stream-stream interaction, as predicted by Goldstein and Jokipii (1977).

While the predominant portion of dissipation occurs at and immediately downstream of the shock, effects far upstream and downstream of the shock can be caused by energetic particle propagation at speeds much greater than the shock. Such energetic electrons (Potter, 1981) and ions (Scholer et al., 1983; Reinhard et al., 1983) and locally generated ion acoustic waves (Kennel et al., 1982) and low frequency whistler mode waves (Tsurutani et al., 1983a; Russell et al., 1983) have been detected in association with flare-generated shock. Both the particles and waves can be viewed as forms of long distance energy dissipation by the shock.

## SHOCK ACCELERATION

In the second process, that of interplanetary acceleration, the charged particles may be energized by a combination of gradient-drifting along the shock surface in the direction of the interplanetary electric field (Hudson, 1965; Chen and Armstrong, 1975; Pesses, 1983) and compression from particle reflection between upstream and downstream cyclotron resonant waves and also between upstream waves and the shock itself (Fisk and Lee, 1980; Lee, 1983). The relative contribution of the above mechanisms is believed to depend on the shock geometry and other shock parameters, such as the upstream plasma beta ( $8\pi \Sigma n k T / B^2$ ) and shock Mach number. Figure 9-6 illustrates the separation of Energetic Storm Particle (ESP) events with quasi-parallel shocks and particle "spikes" with quasi-perpendicular shocks, in good agreement with both theories.

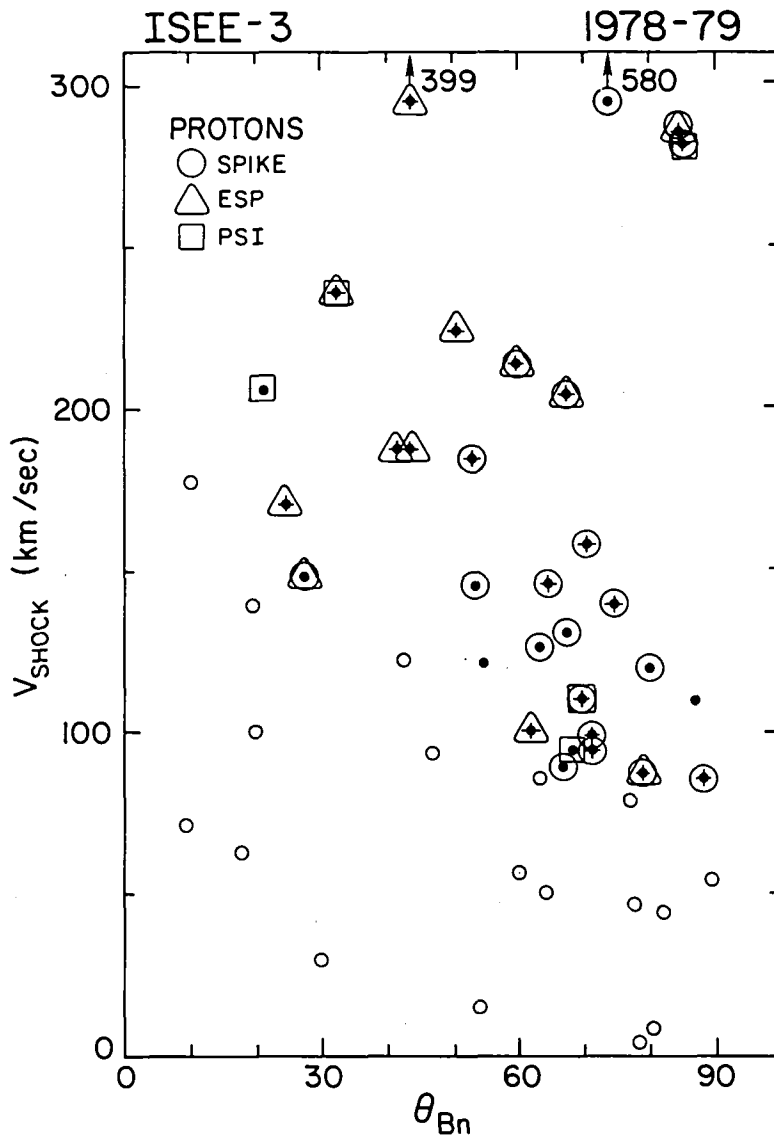


Figure 9-6. Proton event type and intensity as a function of shock velocity relative to the upstream plasma ( $V_{\text{shock}}$ -ordinate) and of shock normal angle ( $\theta_{Bn}$ -abscissa). The types of proton events are: a proton "spike" coincident with shock, an Energetic Storm Particle event, a Post-Shock Intensification, or no detectable change. The symbols for each are indicated in the legend. The particle flux enhancement above ambient background is indicated by the shading: an open circle indicates 0-20% enhancement, a solid dot, 20-200%, and a star, >200%. The distribution of points in the scatter plot indicate that the types of proton events are nicely separated by shock-normal angles. Most of the spike events (large circles) are associated with the shock-normal angles greater than  $50^\circ$  or quasi-perpendicular shocks. The majority of ESP events (triangles) are associated with shocks with  $\theta_{Bn} < 60^\circ$ . There also appears to be a minimum shock velocity that is necessary before sufficient acceleration can take place. This value depends on  $\theta_{Bn}$ . The figure is taken from Tsurutani and Lin, 1983.

A somewhat subtle portion of the shock acceleration question, basic to its resolution, is the complex problem of the source of the "seed" particle population. The relative contributions of heated "downstream" plasma and of remnant solar flare particles is unknown. Additionally, both the shock normal geometry and the plasma beta change with increasing radial distance from the Sun, causing continual evolution of the shock and its associated acceleration efficiency in both distance and time. The effects of the above on the overall acceleration process are inadequately modeled at present.

In major solar flares, such as the August 1972 event, the flux of protons associated with the shock can be as great as the flare particles that arrive at Earth. Thus, this interplanetary acceleration process is fundamental to our understanding of solar flare effects on the terrestrial environment.

### FORBUSH DECREASES

Cosmic ray decreases associated with the passage of flare-associated shocks are well established. Reflection of the energetic particles by the compressed magnetic fields at and behind the shock (Parker, 1961) or by particle drifts in the field gradients (Barouch and Burlaga, 1975) are thought to sweep out the cosmic rays in the path of the high speed stream. Series of shocks have been shown to lead to a succession of decreases in the cosmic-ray flux (McDonald et al., 1981). An example of a sequence of four shocks and four Forbush decreases is shown in Figure 9-7. At this time our understanding of Forbush decreases is only qualitative.

Both the immediate 5-10% cosmic ray decreases associated with a large flare and the combined effect of many flares resulting in the eleven year solar cycle modulation, have significant effects on the terrestrial environment. Cosmic rays can deposit their energy deep in the atmosphere and are the dominant source of ionization in the region below 30 km. Modulation of these particles can have major effects on the local chemistry and on atmospheric electricity. This is discussed in Section V of this Chapter.

### GEOMAGNETIC ACTIVITY

Transfer of energy from the solar flare ejected plasma and fields to the Earth's magnetosphere occurs in two well-established ways. Energy is transferred through solar wind ram pressure changes (Chapman, 1918; Siscoe et al., 1968) and through reconnection between the interplanetary magnetic field and the Earth's field (Fairfield and Cahill, 1966; Arnoldy, 1971; Tsurutani and Meng, 1972; Baker et al., 1983). The former leads to SIs and the latter to magnetic storms. Other solar wind features which have been cited as being important to geomagnetic activity are pressure variations, variations in the IMF (e.g., Garrett et al., 1974) and other causes of "viscous" interaction (Gary and Eastman, 1979; Tsurutani and Thorne, 1982; Gendrin, 1983), and direct plasma injection (Lemaire, 1977; Heikkila, 1982).

Ram pressure increases of 2 to 20 times ambient are caused by the combined velocity and density jumps across fast forward shocks and also by density jumps across tangential discontinuities. The impact of the pressure pulses on the magnetosphere, if sufficiently large, are recorded as impulsive increases of the Earth's field, or Sudden Impulses (SIs). Such compressions of the day-side magnetosphere have important effects on the motion and stability of energetic trapped particles, as will be discussed in the next section.

Although it is now known that at least 90% of all large SIs are caused by interplanetary shocks (E. Smith, 1982) it is generally not known why some SIs are followed by magnetic storms and

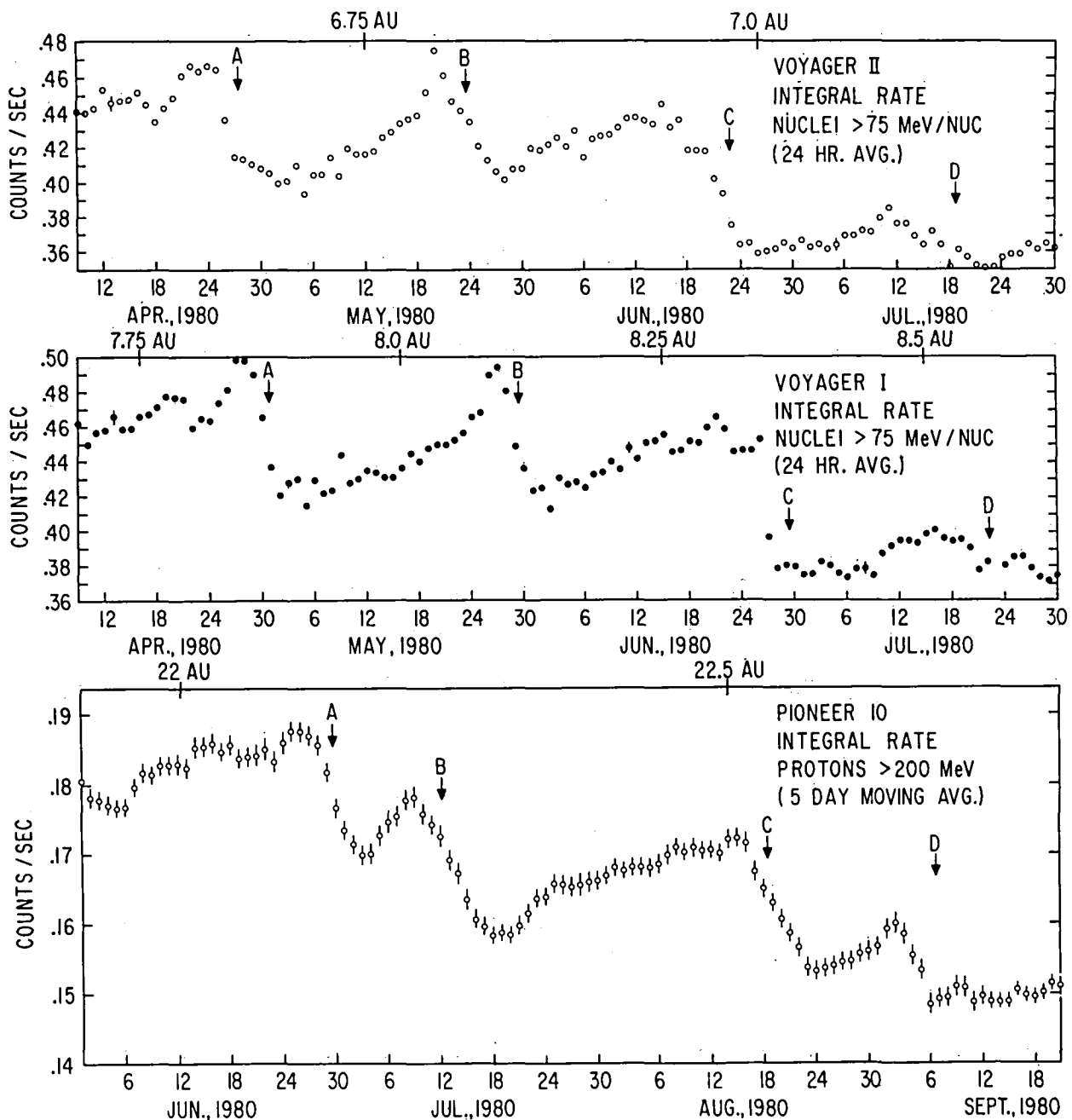


Figure 9-7. A series of four Forbush decreases associated with the passage of interplanetary shocks. Voyagers 1, 2 and Pioneer 10 were at 8, 7, and 22 AU, respectively. The cosmic ray flux decreases with each successive shock, effectively causing a long-term modulation-effect. The figure is taken from McDonald et al., 1981.

others are not (Akasofu, 1977). Even when the solar flare high speed stream does cause a magnetic storm, the length of the storm's initial phase (positive Dst) is highly variable, lasting from minutes to many hours.

Major progress on this topic has been made by the massive study of the recent Coordinated Data Analysis Workshop-6 (CDAW-6). For one flare-associated event, on March 22, 1979, it was determined that the southward interplanetary magnetic fields which were the cause of the two major substorms were related to the magnetic field structure adjacent to a heliospheric current sheet (Tsurutani et al., 1983b). The current sheet was embedded in the shocked solar wind plasma ahead of the driver gas. The current sheet was both highly inclined and tilted relative to the ecliptic plane, causing the large north-south components of the magnetic fields on both hemispheres of the heliosphere. A schematic of the three-dimensional structure is given in Figure 9-8. The causes of this kinked, inclined current sheet and its evolution with distance from the Sun is not currently understood.

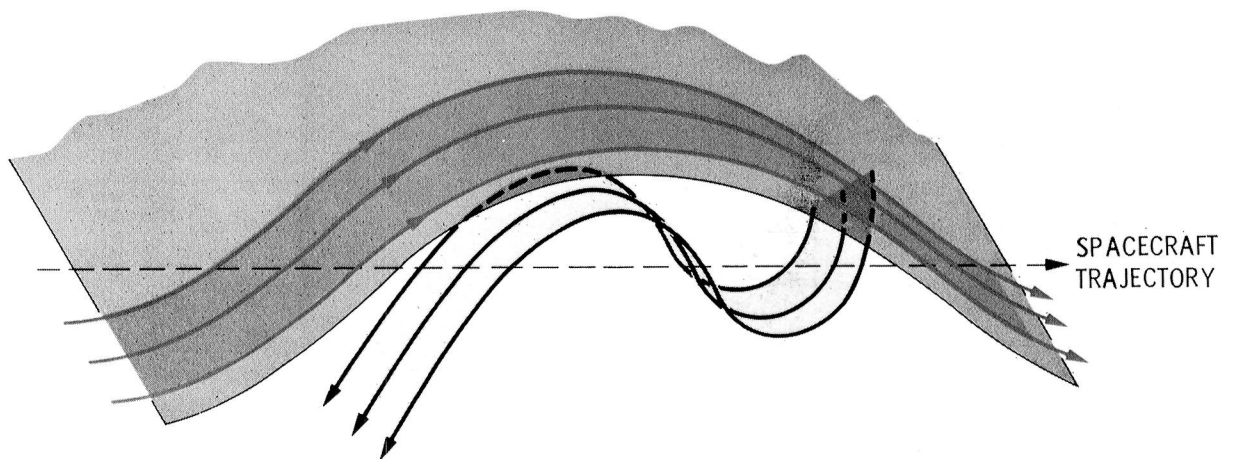


Figure 9-8. A kinky, inclined heliospheric current sheet and neighboring magnetic fields. The kink in the current sheet is even more severe than pictured; the normals to the sheet lie in the ecliptic plane. The current sheet is also tilted up out of the ecliptic allowing for the unusual magnetic field structure in the Southern Hemisphere. Geomagnetic activity at Earth is associated with the passage of two southward magnetic field intervals in the southern heliospheric hemisphere. Substorm intensification occurs when the second current sheet crossing impinges on the magnetosphere., providing a ram pressure increase of six to ten times ambient.



It was further found that large ram pressure increases caused substorm intensifications, a feature which has been noted, but not emphasized, in previous analyses. There are several possible interpretations to these results. Pressure increases (decreases) may only act as a trigger to release energy which is already stored in the magnetotail (Schieldge and Siscoe, 1970). This explanation has been used to explain the near-instantaneous onsets of substorms with SIs. Another possibility is that there may be additional energy transfer into the magnetosphere, either via enhanced reconnection or viscous interaction at the boundary layers. It appears that the latter suggestions are more likely, as the enhancement of substorm energy (AE) occurred during intervals of southward interplanetary magnetic fields.

## OUTSTANDING QUESTIONS

The current task for the next several years is to understand the above relationships better so that quantitative models can be formulated and utilized.

- An energy budget for high speed stream dissipation needs to be developed. The solar wind temperature dependence as a function of heliocentric distance should be modeled.
- The particle acceleration processes need to be developed to a point where the flux and spectrum can be predicted, given the properties of the shock and upstream conditions.
- A quantitative model of Forbush decreases is required. A three-dimensional quantitative model of cosmic ray decreases as a function of the property of the shocks should be formulated, using multiple spacecraft observations of the shocks and cosmic rays as a basis. The eventual goal would be the modeling of cosmic ray decreases everywhere in the heliosphere, given the characteristics of the shock, the upstream plasma and field and the location of the point of observation relative to the flare site.
- It is not known how typical the CDAW-6 results are. More cases studies are needed to be able to determine if kinked, inclined heliospheric current sheets, magnetic bubbles or magnetic tongues (Gold, 1962) are the dominant causes of geomagnetic activity. The effects of pressure changes in the solar wind on magnetic reconnection and/or viscous interaction need to be understood better, both from observational and theoretical standpoints. MHD models of plasma flow from high speed stream sources close to the heliospheric current sheet are needed to understand the evolution and spatial dependence of the magnetic structure ahead of the driver gas.

## IV. IMPACTS ON THE MAGNETOSPHERE

Geomagnetic activity is a consequence of solar activity; it is therefore appropriate to consider the specific mechanisms by which the various manifestations of solar activity can affect the magnetosphere. Solar flares are one such manifestation. Other manifestations include sunspots, solar magnetic-polarity reversals, and (perhaps) the ion-composition of the solar wind. While several of these manifestations of solar activity are likely to have magnetospheric effects, this section is concerned specifically with evaluating the effects of solar flares.

Solar flares can affect the magnetosphere directly by supplying energetic charged particles and indirectly by modifying the interplanetary medium. The interplanetary medium can be modified

through flare-induced changes in solar-wind velocity, density, and ion-composition; through consequent changes in magnitude and direction of the interplanetary magnetic field; and through the formation of shocks, discontinuities, and regions of enhanced plasma turbulence in interplanetary space (see Section III).

## ENERGETIC PARTICLES

The access of solar-flare protons to synchronous altitude (equatorial geocentric distance: 6.6 Earth radii) in the magnetosphere is well documented (e.g., Lanzerotti, 1968; Paulikas and Blake, 1969). Such protons of energy  $\sim 1\text{--}70$  MeV are not necessarily trapped in the Earth's magnetic field, since their trajectories are typically such as to exit the magnetosphere in the absence of scattering. However, an abrupt temporal variation of the magnetospheric B field can easily lead to the geomagnetic trapping of some solar-flare protons (perhaps those in a favored sector of longitude at the time of the magnetic impulse). Once trapped, such protons become a minor part of the radiation belts and are subject to the usual transport and loss processes that apply to radiation-belt protons (e.g., Schulz and Lanzerotti, 1974). These remarks concerning solar-flare protons apply equally to shock-accelerated particles and to the heavier ions ejected by solar flares.

Energetic solar-flare particles over the polar cap and deep within the magnetosphere have been studied extensively as tracers providing information on the configuration of the geomagnetic field and, in particular, on the degree to which geomagnetic and interplanetary magnetic field lines are connected. The uniformity of electron intensities in the domain of open field lines near the poles suggests that the magnetotail is filled quite uniformly with solar flare electrons traveling along field lines directly connected to the interplanetary field rather than through diffusion (Anderson and Lin, 1969). Thus, it is generally concluded that the direct access of solar electrons to the magnetosphere strongly supports an open magnetospheric model.

Recently solar flare neutrons of energy 50-600 MeV have been observed in space aboard the Solar Maximum Mission spacecraft (Chupp et al., 1982). The measurements suggest that approximately  $10^{28}$  neutrons/sr above 50 MeV were emitted by the Sun and this, in turn, requires the rapid acceleration of protons to GeV energies during the impulsive phase of the flare. These neutrons are important in flare studies but (at flux levels of the order of  $4 \times 10^{-2}$  neutrons/cm<sup>2</sup>-sec) do not impact magnetospheric dynamics.

## SOLAR WIND EFFECTS

The plasma emitted by a solar flare (see Section III) is likely to travel faster (and be richer in helium) than the ambient solar wind (e.g., Hundhausen, 1972). The increased helium abundance may eventually express itself in the Earth's ring current. The increased plasma velocity produces a shock front as the flare plasma overtakes the ambient solar wind plasma. The Earth's magnetosphere will receive an impulse as the shock front encounters the magnetopause. This point has been well established and the correlation between shocks and sudden impulses is now thought to be about 90%. Such an impulse has both adiabatic and non-adiabatic consequences for geomagnetically trapped particles (e.g., Schulz and Lanzerotti, 1974). The adiabatic consequence produces a reversible energization of such particles. The non-adiabatic consequence contributes to radial diffusion.

Propagation of the flare-produced shock front through interplanetary space changes the magnetic-field configuration there (see Section III). There is a notable compression of the interplanetary magnetic field where the flare plasma overtakes the ambient solar wind plasma. The result is an

outward-propagating magnetic mirror and enhanced particle scattering that may partially shield the inner solar system from incident cosmic rays. Such a mirror might also shield the Earth's magnetosphere against the 1-5 MeV electrons from Jupiter that are normally present (e.g., Krimigis et al., 1975) in the interplanetary flux tube that connects the Sun to Jupiter's magnetosphere. A permanent change of solar wind velocity in a longitude sector emanating from the flare site would shift the phase angle required for the Earth and Jupiter to be in the same flux tube. Baker et al. (1979) have shown that enhancements of the radiation-belt electron intensity ( $E > 1$  MeV) at synchronous altitude tend to follow by about two days the appearances of Jovian electrons just outside the Earth's magnetosphere. They have inferred from this that Jupiter may at times be an important source of such high-energy electrons for the Earth's outer radiation belt.

Cross correlation analyses between various interplanetary parameters and geomagnetic activity indices using large data sets indicate that the main relationship is with southward interplanetary magnetic fields (IMF) and to a lesser degree with the solar wind ram pressure. Although observations following sudden impulses (Arnoldy and Moore, 1982) indicate that convection within the magnetosphere is increased during these periods, as indicated by geostationary satellite measurements, the main phase of a storm (development of the ring current and substantial negative Dst) may or may not follow. However, numerous studies have indicated a high correlation coefficient between geomagnetic activity and the IMF  $B_z$  component, giving indirect evidence that dayside magnetic "reconnection" is responsible for the coupling between the solar wind and the magnetosphere. The relationship between  $B_z$  and geomagnetic activity is a general one and has not been developed separately for magnetic storms, although it is generally assumed that such a relationship exists.

Flare-produced changes in solar wind pressure and/or in the direction of the interplanetary B field in the vicinity of the Earth's magnetosphere are likely to alter the rate of solar wind energy coupling into the magnetosphere (see Section III). Consequent changes in the strength of the Earth's magnetospheric electric field contribute to the radial diffusion of geomagnetically trapped particles (e.g., Schulz and Lanzerotti, 1974) as well as to a host of other magnetospheric phenomena. These phenomena include changes in magnetospheric convection velocities, in auroral activity and location, and in ionospheric currents, as well as erosion (or replenishment) of the plasmasphere. Erosion of the plasmasphere is normally accompanied by enhancement of the ring current, since both phenomena are contingent on an increase in the magnetospheric electric field. It is thus presumably through "reconnection" that solar flares are related to magnetic storms and substorms and the various other manifestations of geomagnetic activity.

## OUTSTANDING QUESTIONS

"Reconnection" is defined by Vasyliunas (1975) as the process whereby plasma flows across the boundary that separates magnetic-field lines of different topology. In the present context the different topologies are interplanetary and magnetospheric, and the effect of "reconnection" is to permit the interplanetary electric field to have a component tangential to the magnetopause. This is a macroscopic description of the process. One of the major conceptual uncertainties in magnetospheric physics involves identification of the microscopic plasma process (if any) that allows reconnection to occur. Other aspects of magnetospheric physics seem to be broadly understood in concept if not in quantitative detail. These include charged-particle transport in the magnetosphere; conversion of solar-wind energy to the production of energetic particles; and various aspects of storm/substorm dynamics, including auroral phenomena. The deposition of energy by 10-100 MeV protons into the high latitude and polar cap ionosphere results in Polar Cap Absorption (PCA) events. These events have been studied in the past primarily in terms of radio-wave

absorption (see Section V). With a better appreciation now of the close coupling between the ionosphere and the magnetosphere, the effect of the greatly enhanced conductivity produced in the ionosphere by these events on the topology and dynamics of the magnetosphere is a topic for future study.

## V. IMPACTS ON THE IONOSPHERE AND ATMOSPHERE

The aspects of solar flares which directly affect the Earth's atmosphere and ionosphere are: (1) enhanced X-ray and EUV emissions which arrive at the Earth minutes after flare commencement; (2) protons, alpha particles and electrons which arrive several minutes to several hours later (depending on the solar longitude of the flare region) and are deposited in the Earth's polar caps; and (3) energized solar wind plasma which arrives at the Earth after the solar particles and causes magnetospheric storms. In this section we describe the mechanisms by which the solar X-ray, EUV and particle emissions induce changes in the Earth's atmosphere. We then discuss the nature and extent of the disturbances associated with these events and the uncertainties in our current understanding.

### X-RAY AND EUV EFFECTS

The EUV and X-ray flux enhancements in a solar flare are highly variable. The flux generally decreases with increasing wavelength but different wavelengths do not necessarily increase at the same rate or at the same time. The flux can also vary substantially from flare to flare. For example, the Sun's total X-ray emission can rise by several orders of magnitude in the range between two and ten Angstroms. In the same flare, the Lyman  $\alpha$  radiation (1216 Å) might increase by only a few percent. The enhanced EUV and X-rays reach the Earth minutes after flare commencement and immediately affect the D, E, and F regions of the entire sunlit ionosphere. The EUV radiation is mostly absorbed in the E and F regions while the X-rays penetrate deeply into the D region below 100 km. The immediate ionospheric effect of the enhanced radiation is an increase in the electron densities. The ionospheric changes are persistent at higher altitudes, where transport has a smoothing effect and ion recombination times are long. At lower altitudes, in the D and E regions, the time scale of electron density perturbations last roughly as long as the 10-30 minute flare perturbations.

In the quiet daytime ionosphere, the major ion in the upper ionosphere is  $O^+$ . In the lower ionosphere the major ions produced are  $N_2^+$  and  $O_2^+$  which undergo chemical reactions to produce  $NO^+$ . During a solar flare, however, the dominant ion in the E and upper D region can be  $O_2^+$  instead of  $NO^+$ . Thus solar flares can change the basic pathways of the ion chemistry in the ionosphere. Since charge neutrality demands that an equal number of electrons be produced as positive ions, the ionospheric electron density changes to reflect the flare-induced ion changes. Figure 9-9 compares typical electron density profiles for a large flare with those typical of the quiet daytime ionosphere.

Extensive studies have been done to produce models of flare X-ray and EUV enhancements and the resulting ionization (Donnelly, 1976; Donnelly et al., 1976). Detailed time-dependent models of the chemical and resulting electron density changes have provided a detailed picture of the disturbance (Oran et al., 1976; Oran and Young, 1977; Mariska and Oran, 1981). Enhanced electron densities predicted by such studies have been shown to reproduce the anomalous radio propagation effects such as those observed in the intense solar flare of August 1972.

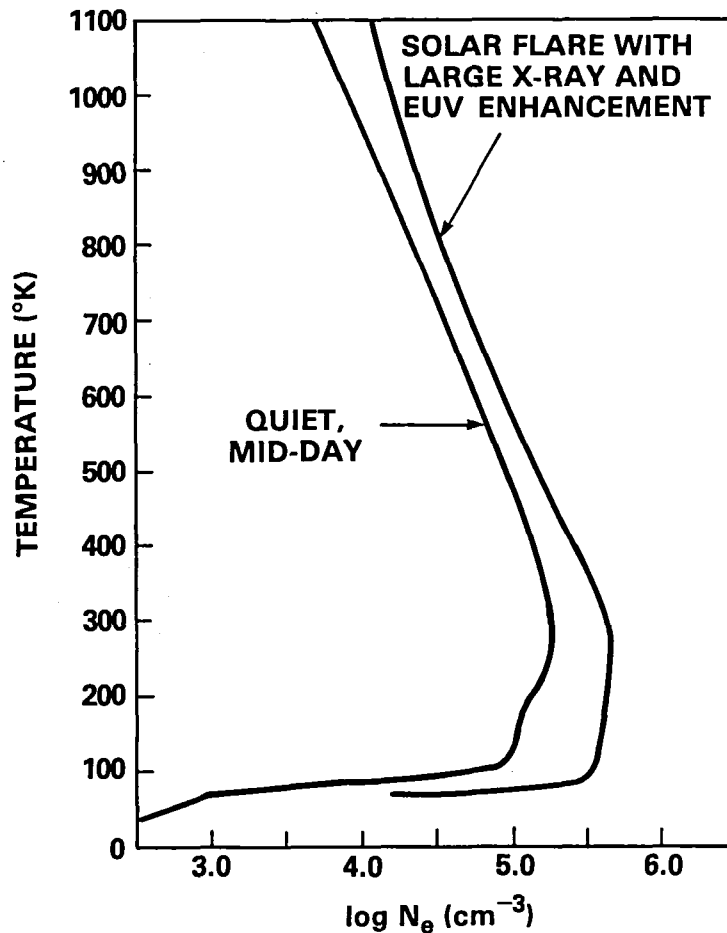


Figure 9-9. Comparison of typical electron densities in the quiet daytime ionosphere with those produced by the enhanced X-ray and EUV emissions of a large solar flare.

In addition to the direct chemical changes, EUV and X-ray enhancements can cause dynamical effects. For example, large enhancements in conductivity due to increased ionization have been found to excite MHD waves in the magnetosphere (Rosenberg et al., 1981). In this coupling of solar radiation to these MHD waves, the ionosphere is used as an intermediary and can trigger magnetospheric processes.

## SOLAR PARTICLE EFFECTS

Large flares often produce intense fluxes consisting mostly of protons and alpha particles. These particles have energies that extend to several hundred MeV. The only source which is more energetic is that of the galactic cosmic rays. Electrons with energies up to a few MeV are also present at times. When the energetic particles, called a Solar Particle Event (SPE), reach the Earth and penetrate the atmosphere, the main ionization is often produced below 60 km and can last from several hours to several days. As shown in Figure 9-10, the solar proton ionization rate in the August 1972 SPE peaked near 40-45 km and exceeded the background ionization produced by cosmic rays by over four orders of magnitude. The entire north and south polar cap regions extending down to latitudes as low as 45 degrees were affected.

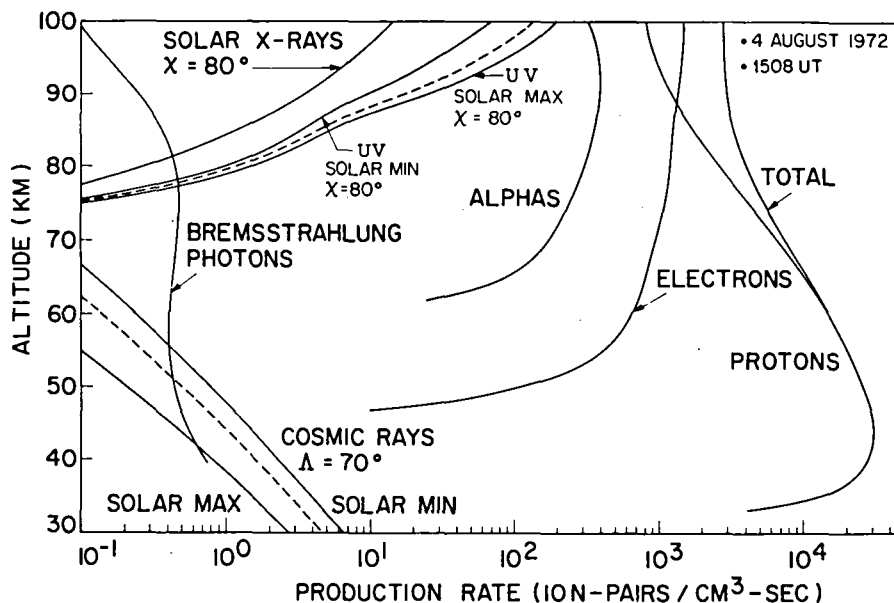


Figure 9-10. Ion production rate in the lower ionosphere produced by the large SPE of 4 August 1972. For comparison the normal production rates due to solar UV, X-rays and cosmic rays are shown (from Reagan and Watt, 1976 reprinted with permission from the J. Geophys. Res.).

Efforts have been made to model the complex ion chemistry changes in the mesosphere during the large SPEs on November 2, 1969 (Swider et al., 1978) and August 4, 1972 (Reagan and Watt, 1976; Reagan et al., 1981). Solar particle spectra measured on satellites, balloons and rockets are used with atmospheric deposition models to obtain ion production rates as a function of altitude. The ionization rates obtained are believed to be accurate to about 25%, which is considerably better than our knowledge of many of the chemical reactions in the D region. The ionization rates are used with chemistry models that deal with the greater than 200 reactions that are important in the D region. In contrast with the E and F regions, negative ions are extremely important in the D region. The SPE greatly affects the ion chemistry paths and even alters key elements of the neutral chemistry, such as atomic oxygen, that influence the final electron density. The large ion densities, in excess of  $10^5$  ions/cm<sup>3</sup>, produced below 70 km by an SPE significantly affect both the amplitude and phase of transpolar VLF signals used for navigation and military communications. Electron densities in excess of  $10^5$  cm<sup>-3</sup>, which create blackout at high frequency (HF), were both measured and modeled at altitudes near 60 km in the August 1972 SPE. Such high electron densities normally exist only near the F-region peak.

Finally, large solar flares can both increase and reduce the ion concentrations in the lowest regions of the atmosphere. In many SPEs the proton energy spectrum extends above 1 GeV and can penetrate to the Earth's surface. During such occasions, known as ground level events (GLE), the ionization from altitudes of 60 km to the surface, which ordinarily is dominated by galactic cosmic rays, is increased. When the magnetic field in the flare-enhanced solar wind interacts with the interplanetary magnetic field, a portion of the galactic cosmic ray flux is shielded from the Earth. During these Forbush decreases, the ionization near sea level is reduced by 5% to 30%. Although the absolute intensities and changes involved in both the GLE and Forbush decrease are relatively small, the cosmic rays are the dominant ionization source in the lower

ionosphere and to altitudes as low as a few meters above ground level. Thus, cosmic rays play an important role in determining the ionospheric potential that is part of the global atmospheric electric circuit.

## MAGNETOSPHERIC STORM EFFECTS

Large solar flares typically produce an enhanced solar wind that interacts with the Earth's magnetosphere one to two days later. This interaction usually produces a major magnetic storm that affects the Earth from mid-to-high latitudes. The resulting electron and proton precipitation from the Earth's radiation belts greatly enhances the ionization and conductivity of the ionosphere. In the auroral E region the intense currents that flow in the electrojet produce significant Joule heating of the local ions and electrons. The large  $E \times B$  forces drive the ions and electrons into large-scale motion. The high-latitude ionospheric disturbances are transported to low latitudes, even to the equator, over the period of a few days in a manner that is not totally understood. In addition to local ionization and heating enhancements, the F region is also subject to bulk transport of ionization from below by mechanical forces in the form of gravity waves; the result is an upheaving of the ionosphere. Dramatic ionization increases can occur in the nighttime D region where electron densities at 90 km can exceed  $10^5 \text{ cm}^{-3}$ . The enhanced ionization during a storm can extend to magnetic latitudes as low as 45 degrees. At mid-latitudes, the subsequent electron drizzle from the enhanced radiation belts may influence the mesospheric ionization for up to 10 days (Reagan, 1977) or until another major storm occurs.

## EFFECTS ON RADIO WAVE PROPAGATION

Solar flare effects in the ionosphere have been primarily observed by their effects on radio wave propagation, which is altered extensively by changes in the ionization profiles. Table 9-1 lists the types of events described above and the aspects of these events important for radio wave propagation. The general class of phenomena called sudden ionospheric disturbances (SIDs) collectively describes a number of irregular, impulsive events which affect the intensity, frequency and phase of radio signals. These effects are caused by the changes in ionization with altitude and time as well as by changes in the total electron content through which the signal propagates. More specifically, the long-range transpolar transmissions at very low frequencies (VLF), used for navigation and military communications, may be severely disturbed. Also, civil and military communications at high frequencies (HF) may suffer "black-out" for many hours in a large SPE. For example, during the intense SPEs of August 1972, HF black-out conditions at high latitudes persisted for several days.

TABLE 9-1  
DISTURBANCES TO IONOSPHERIC PROPAGATION

EVENT	DISTURBANCE	IONOSPHERIC EFFECT	PROPAGATION EFFECT	DISTURBANCE DURATION
SOLAR-FLARE (X-RAYS/EUV)	SID	INCREASED D- AND E-REGION ELECTRON DENSITY BY DAY	INCREASED ABSORPTION: LOSS OF SIGNAL AT HF	10-30 MINUTES
SOLAR PARTICLE EVENT	PCA	HIGH D-REGION ELECTRON AND ION DENSITIES OVER POLAR CAPS	LOSS OF POLAR COMMUNICATIONS AT HF (BLACKOUT); SERIOUS EFFECTS AT ELF AND VLF	FEW TO SEVERAL DAYS
MAGNETIC STORM AND SUBSTORMS	STORM EFFECT	INCREASED D- AND E-REGION ELECTRON DENSITY AT MID AND HIGH LATITUDE  DECREASED F-REGION ELECTRON DENSITY	AMPLITUDE ATTENUATION AND PHASE SHIFTS AT VLF: FADEOUTS AT HF	SEVERAL HOURS TO SEVERAL DAYS

## NEUTRAL ATMOSPHERE EFFECTS

The energy deposited by solar particles can also change the neutral chemistry of the upper atmosphere. The most significant changes occur in the concentrations of key minor constituents such as atomic oxygen, ozone, hydroxyl radical, hydrogen dioxide and the nitrogen oxides. These three groups of minor constituents known as odd-oxygen, odd-hydrogen, and odd-nitrogen, play an extremely important role in the chemical balance of the atmosphere. For example, the solar ultraviolet absorption by ozone is the principal heating mechanism in the middle atmosphere. Changes in the heating rates of the mesosphere and stratosphere affect the dynamics of the atmosphere and may affect the Earth's albedo. Whether such changes are sufficient to produce changes in the troposphere where the Earth's weather system is formed is not known and is a research area of major interest and controversy. Although it now appears that solar disturbance effects may not be the major cause of stratospheric warnings, the issue is still open.

Changes in the upper mesosphere ozone concentration have now been reported in the SPEs of November 2, 1969, August 4, 1972 and September 12, 1982. In moderate SPEs where the proton deposition is principally above 40 km altitude, the ozone depletion is due to the creation of odd-hydrogen compounds that catalytically reduce the odd-oxygen species ( $O$  and  $O_3$ ). The ozone reduction is on the time scale of many hours to a few days and leads to a decrease in atmospheric heating rates. Even moderate SPEs may create such an effect. It has been estimated that the reduction in ozone in the August 4, 1972 SPE resulted in a 4K temperature decrease at altitudes near 40 km several days later (Reagan et al., 1981). In SPEs where the proton spectra is very energetic and intense, the primary energy deposition can occur in the stratosphere at altitudes near 40 km as shown in Figure 9-10. At these altitudes the level of compounds ( $NO$ ,  $NO_2$ ) is enhanced. These levels persist and can react catalytically with atomic oxygen and ozone to cause a reduction of concentration. During the August 1972 SPE, the upper portion of the Earth's ozone layer was depleted by this effect. A polar ozone cavity defined in latitude and longitude by the solar proton extent was created in which the enhanced  $NO$  concentrations and the corresponding  $O_3$  reductions persisted for as long as 50 days (Reagan et al., 1981). The oval-shaped cavity rotated around the polar regions with the prevailing winds. Within the cavity the ozone column content was estimated to be reduced by  $\sim 2\%$ , as shown in Figure 9-11.

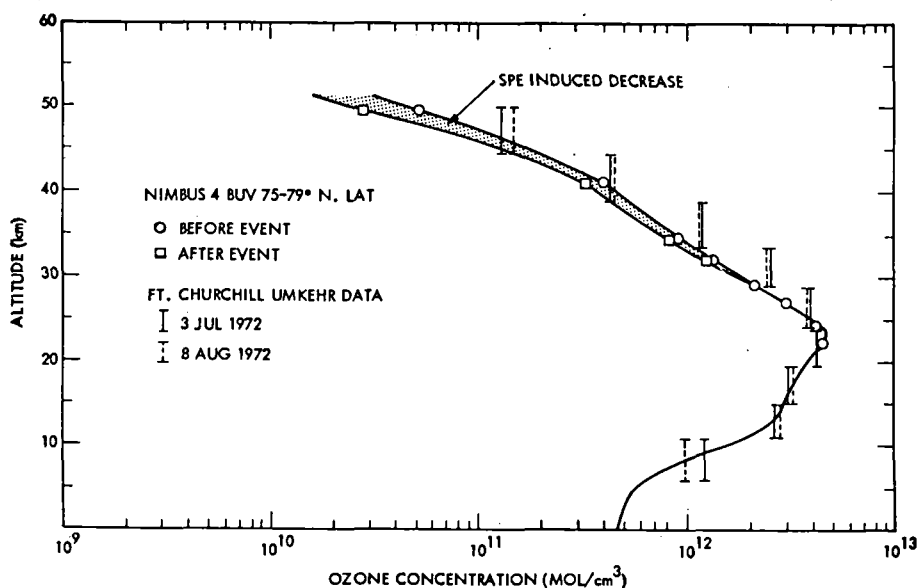


Figure 9-11. Effects of the 4 August 1972 SPE on the high latitude stratospheric ozone layer. The ozone column content was estimated to be reduced by 2 percent due to the SPE (from Reagan et al., 1981 reprinted with permission from the J. Geophys. Res.).



## ATMOSPHERIC ELECTRICITY EFFECTS

Solar particle precipitation also greatly enhances the conductivity of the lower atmosphere. Conductivities in the August 1972 SPE were increased by factors of 200 near 40 km. In addition, the currents carried by the solar protons and electrons greatly exceeded the normal air-Earth current that maintains the global atmospheric electric circuit driven by worldwide thunderstorms (Reagan et al., 1983). The result of these two effects is that the vertical component of the Earth's electric field was shorted-out down to altitudes as low as 20 km, as shown in Figure 9-12. Confirmation of this fact was provided by balloon-borne electric field measurements. The large Forbush decrease associated with the SPE resulted in a decrease in conductivity between ~12 km and the ground and may have been more important to the global circuit since most of the resistance in the circuit occurs at these altitudes.

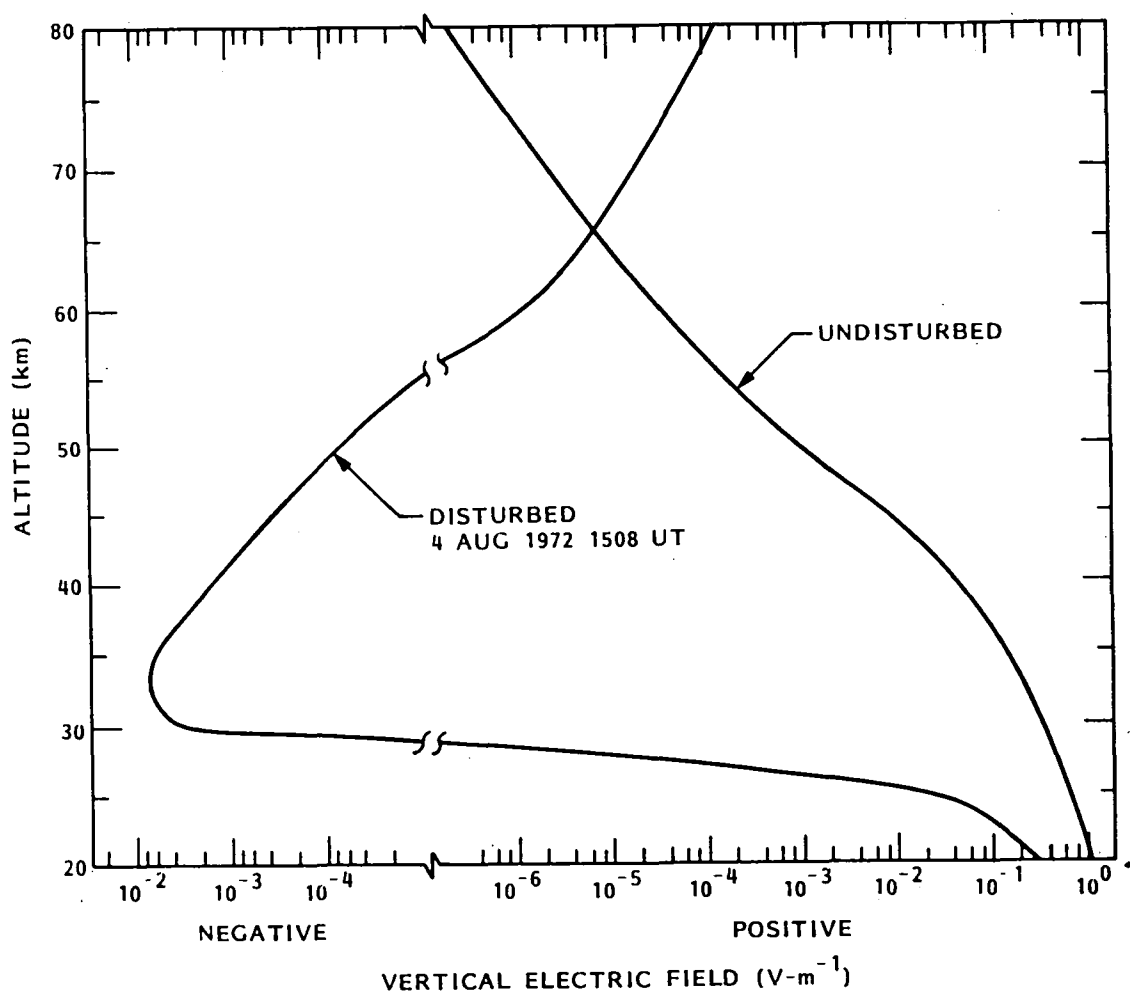


Figure 9-12. Effects of the 4 August 1972 SPE on the vertical component of the atmospheric electric field at high latitudes. The electric field was shorted-out by the SPE between altitudes of 30 and 55 km (from Reagan et al., 1983 reprinted with permission from the J. Geophys. Res.)

The density of the thermosphere varies with magnetic storms and over the solar cycle. A large magnetic storm heaves the atmosphere upwards and the neutral density increases above 100 km altitude can be large. Such density increases produce substantial drag on low-altitude satellites and can result in premature re-entry from orbit, as experienced by Skylab in 1979.

## OUTSTANDING QUESTIONS

The major uncertainties in our understanding of the effects of solar disturbances on the ionosphere and atmosphere deal with the global inputs and transport processes. The global issue involves a knowledge of the disturbance sources on a worldwide basis. For example, no model exists for the input produced by magnetospheric electron precipitation. The transport issue deals with the atmosphere, ionosphere, and magnetosphere as a coupled system. For example, transport of high-latitude ionospheric disturbances to low latitudes is not well understood. Other more specific questions concern the extent to which magnetospheric storm and substorm particle precipitation control the nighttime D-region ionization at mid- and high-latitudes, and the extent to which energetic electron precipitation may be affecting ozone concentrations in the mesosphere. More speculative issues concern the influence of solar flare and particle events on weather and climate through ozone depletion and changes in atmospheric heating rates, changes in heating rates affecting the Earth's albedo, or changes in electrical characteristics of the atmosphere affecting thunderstorm frequency and electrical current output.

Based on the uncertainties mentioned above, future research with the goal of improving our understanding of the effects of disturbances in the Earth's atmosphere should take the following direction. First, we would like to understand the physics of the Sun well enough to predict the occurrence of flares and their magnitudes. Since this is not likely in the near future, we need better monitoring of the Sun to provide real-time data on X-ray, EUV and solar particle enhancements. In the same sense, we need to know the precipitation into the ionosphere from the magnetosphere on a global scale. Optical and X-ray imaging of the globe from high altitude spacecraft may be our best method of accomplishing this. Given the unperturbed ionospheric background and the solar X-ray and EUV inputs, we would then be in a good position to use the known properties of E- and F-region chemistry to calculate ionospheric electron densities. Given a spectrum of the enhanced X-rays or protons, we can calculate the D-region ionization production rates to high accuracy. The conversion to electron densities is less certain because of the uncertainties in the D-region chemical processes.

Coordinated experimental efforts to investigate the effects of solar disturbances on the Earth's coupled atmospheric system are needed. Such efforts must include simultaneous satellite, balloon, rocket and ground measurements to cover the altitude regime from the Earth's surface to the magnetosphere. Particular attention must be paid to the long-lasting solar proton events. As part of this overall effort should be the development of multidimensional global ionospheric and atmospheric models which can be tested by the experimental results. Optical and X-ray imaging of the globe from high altitude spacecraft may be our best method of accomplishing this.

## VI. IMPACTS ON TECHNOLOGY

Since the early 18th century, when concepts of electricity and magnetism began to be used for practical purposes in the electrical telegraph, effects of the natural environment on technology have been concerns for engineering, science, and commerce (Barlow, 1949; Prescott, 1860). The recognition that solar flares are the cause of geomagnetic disturbances, and thus the source of many of the "natural" disturbances of technology, is a somewhat more recent development. As the "biosphere" in which humans live and place their technological achievements expands to include the space around the Earth, the solar-originated processes which can affect technology have increased both in number and in the nature of the effects. Furthermore, as the sophistication of technology increases, more subtle effects of the Earth's environment, and therefore of solar processes, must be considered when implementing new engineering designs that operate in, or make use of, the natural environment.

Because of the limitation of space, only a few examples of the effects of the solar environment on technology are given. A recent overview is presented in Paulikas and Lanzerotti (1982).

### RADIO COMMUNICATIONS

Commercial telecommunications long ago progressed to underground and undersea cables, microwave links, and satellite paths to increase the bandwidth available and to avoid the propagation problems inherent in long-distance signal transmission via ionosphere reflections. However, ionosphere communications are still of great usefulness to activities in the polar regions (which are often inaccessible to synchronous satellite links), to some developing nations, and to military communications and navigations (see also Section V).

Although the commercial implications of ionospheric variability are presently small, the nuisance aspect of the phenomenon is real. For example, increasing solar activity substantially alters the ionization levels. These altered conditions can be determined reasonably well, given data on the inputs to the ionosphere and upper atmosphere, but predictions of their precise effects for specific transmission paths is difficult. An increase in ionization during increasing sunspot conditions can, for various lengths of time, produce anomalously long-path propagation of signals which are intended to be only local. For example, the distress calls from the crash of a small plane in West Virginia in February 1979 were heard in Orange County, California, causing considerable confusion in the local civil emergency procedures (Emmons, 1979). Such situations as these are not isolated during intervals of solar storminess and are certainly a problem to the people and institutions concerned.

### SCINTILLATIONS OF SATELLITE SIGNALS

The use of transionosphere satellite communications links eliminated the problem of variable ionospheric reflection properties for long-distance signal propagation. However, the small-scale structures of the ionosphere have become important because of the "scintillation" which they produce in the signals. Such scintillations set fundamental limits for satellite communications, radar ranges and directional sensing. The effects of these scintillations in disturbing communications links have extended to satellite frequencies in the GHz range, a frequency range that was not expected to be affected by ionosphere irregularities. The scintillations measured in the INTELSAT network at 4 and 6 GHz are observed to have peak-to-peak amplitudes of  $\sim 1$  to 10 dB and to occur predominantly in the equatorial regions of the Earth (Taur, 1973, 1974, 1976). At 4 GHz, the power spectrum of the scintillations typically falls with frequency as  $\sim f^{-3}$  for frequencies  $> 0.125$  Hz (Taur, 1979).

Some satellite and rocket data are now available that suggest that the Rayleigh-Taylor instability occurring in the current flowing in the "equatorial electrojet" produces the scintillation-causing irregularities (Kelley et al., 1976; Scannapieco and Ossakow, 1976; Chiu and Straus, 1979). A chapter in this document is devoted to the physics of scintillation phenomena.

## CABLE COMMUNICATIONS

Associations between geomagnetic storms and disturbances of cable communication systems have been observed for many years (Germaine, 1940; Winckler et al., 1959; Anderson et al., 1974). The highly variable (in both amplitude and time) geomagnetic field during magnetic storm conditions induces Earth currents, and thus Earth surface potentials, across the length of a cable. The output voltage of the Oban, Scotland, power-fed equipment on the transatlantic cable from Newfoundland showed an excursion greater than 1 kV during a magnetic storm in 1958 (Figure 9-13). Detailed study of the August 4, 1972 communications cable disruption in Illinois revealed them to be a result of large compressions and asymmetric distortions of the Earth's magnetosphere and not to intensifications of the auroral current system. Greatly enhanced currents flowing on the magnetopause are believed to have been the primary cause of induction of the Earth currents (Anderson et al., 1974).

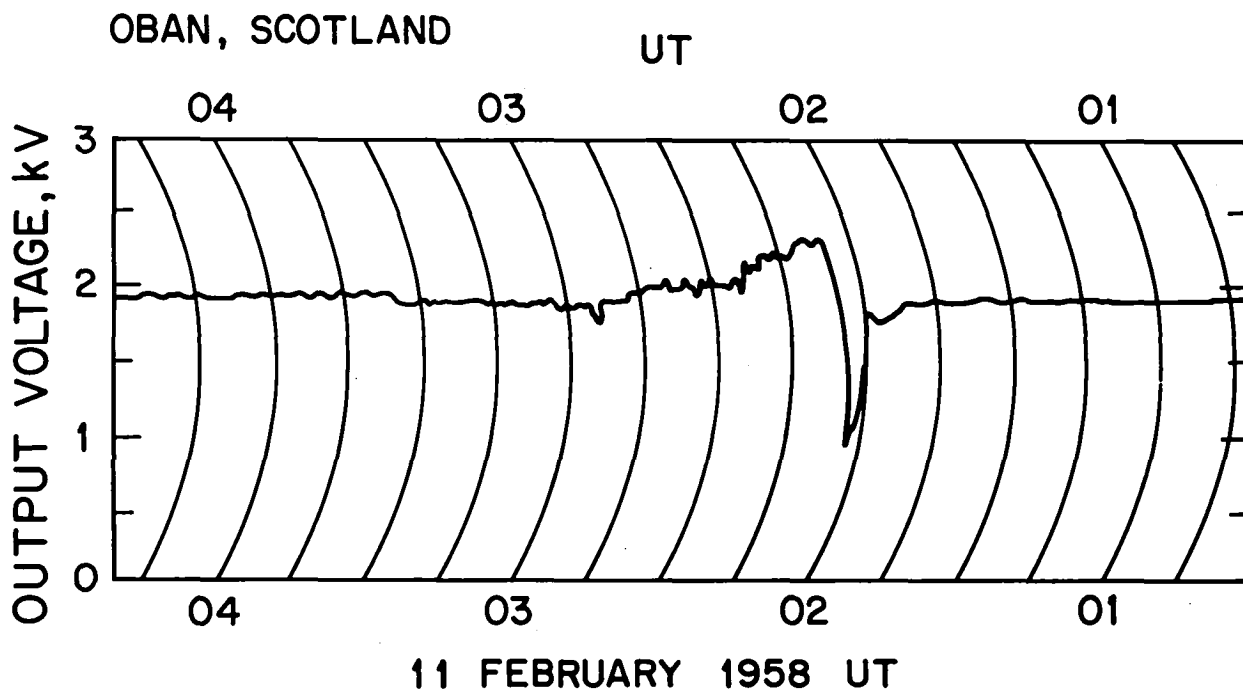


Figure 9-13. Output voltage of the power feed equipment at the Oban, Scotland, end of the Oban-Clareville, Newfoundland, cable. The voltage variation in North America was somewhat larger, leading to a total variation of about 2700V across the cable (from Axe, 1968).

The magnetic field changes observed in the area of the cable disruption were employed in a model calculation, using a three-layer Earth conductivity model, to estimate the potential difference along the route between relay powering stations (Anderson et al., 1974). It was found that the surface potential drop exceeded by at least 10% the Earth-potential design limit for the system. While such disruptions have been infrequent in recent years, principally because of increased system design margins, the uncertainties in our knowledge of the scale sizes of large geomagnetic disturbances imply that we cannot yet be absolutely confident of having eliminated the geomagnetic induction problem from possible disruptions to cable communication systems.

## GROUND-BASED POWER SYSTEMS

As for cable communications, disruptions of power systems by geomagnetic disturbances are well-documented (Albertson et al., 1973; Davidson, 1940; Slothower and Albertson, 1967). For example, during a magnetic storm in 1958 Toronto, Canada, was plunged into a temporary black-out because of the tripping of circuit breakers in an Ontario transformer station (Brooks, 1959).

The geomagnetic currents induced in a power system can produce problems of several different types (e.g., Aspres et al., 1981). On the one hand, the arbitrary differential relay operation in power distribution systems during geomagnetic storms can produce a judgmental problem; systems operators are unsure of whether or not the malfunctioning relay is an indication of an induced current effect in a transformer or a real transformer malfunction. On the other hand, the currents actually induced in the winding of a power transformer can result in half cycle saturation of the transformer core. This saturation can produce fluctuations in the distribution system voltages and intense localized heating of a transformer itself. This local heating can greatly shorten the lifetime of a transformer. In extreme situations, outright failure or fire can result.

## PIPELINE

A pipeline is another type of long conductor which can be affected by induced events caused by geomagnetic activity. There does not appear to be a severe corrosion problem on present-day pipelines from induced currents, provided cathodic protection circuits are operating. Rather, induced currents can interfere with engineering work associated with normal pipeline corrosion surveys.

An area of active study is the effect of auroral-induced events on the Alaska pipeline, which extends for  $\sim 1.3 \times 10^3$  km from the Arctic Sea to Valdez on the southern coast of the state. A comparison of the event induced in the pipeline as measured near Fairbanks (Chena) and the Earth currents measured at Fairbanks (College) is shown in Figure 9-14 (from Campbell and Zimmerman, 1980). Probably the most important effects that can rise from the currents are their impacts on the pipeline monitoring and control electronics. Further, large induced currents of a transient nature can greatly disrupt, or even prevent corrosion survey engineering studies on the pipeline.

## THE SPACE RADIATION ENVIRONMENT AND SPACE SYSTEMS

When communications satellites were first proposed there was no thought that the space environment for these technological developments would be anything but benign. However, the discovery of the Van Allen radiation belts immediately indicated that the particle radiation environment of spacecraft would be of a major concern to satellite designers.

The fluxes of low energy ( $\lesssim 20$  keV) protons and electrons at synchronous orbit are sufficiently intense to produce degradation of spacecraft thermal control surfaces and exterior coatings. This portion of the particle energy spectra also determines whether or not spacecraft surface charging

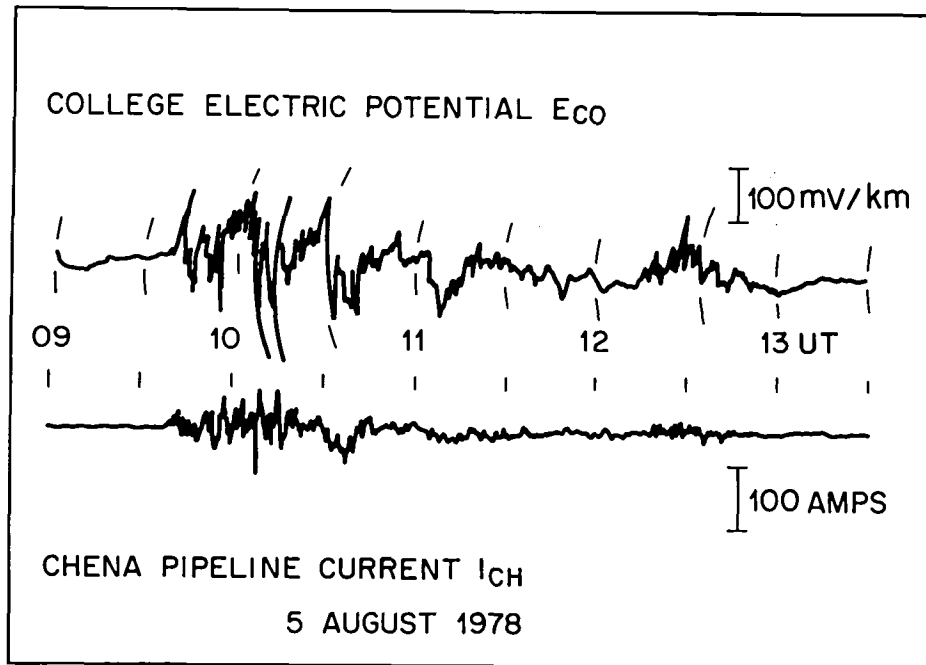


Figure 9-14. Comparison of the Earth currents measured near Fairbanks and the induced currents measured in the Alaskan pipeline at the Chena test site (near Fairbanks) on 5 August 1978 (from Campbell and Zimmerman, 1980.)

to high potentials (DeForest, 1979) will occur at a given time. Higher energy particles, which can penetrate a spacecraft, can produce charging in dielectrics, such as coax cables in the interior of the vehicle.

Solar cells are completely covered to avoid the high specific ionization of trapped, MeV-energy, protons. The trapped electrons then provide the major cause of radiation damage to solar cell arrays and to semiconductor devices deep in a spacecraft interior. A typical solar cell array loses a few percent of the original output power per year of exposure to the space radiation environment (Paulikas, 1979). Step-like decreases can occur from a large enhancement in the synchronous altitude proton fluxes following a large solar flare event such as that in August 1972.

Although spacecraft designers typically want to know the radiation environment to an accuracy of 10%, present predictive capabilities are no better than a factor of two or greater. Order-of-magnitude variations in the electron and proton fluxes are common and frequently unpredictable.

The technological improvements in satellite capabilities through the use of low-power CMOS devices have also resulted in a more severe radiation hazard problem; these devices often do not have the radiation resistance of the higher power TTL logic that they replace. The increasing use of large scale integrated circuit devices requires more accurate models of the trapped radiation and its changes with solar activity, including models for even the esoteric galactic cosmic radiation. For example, it is believed now that heavy nuclei in the galactic cosmic ray flux must be considered as explanations for the upsets seen in highly miniaturized satellite electronic components (Kolasinski et al., 1979). It may be necessary to consider the modulations of the galactic radiation by solar flare shock waves (see Section III) in future designs.

## PEOPLE IN SPACE

The radiation from solar flares and solar flare-induced effects can be harmful or even lethal to humans exposed to it. Extensive monitoring of the Sun was carried out during the Apollo missions to the Moon in order to ensure, insofar as possible, flights during intervals of reasonably low solar activity, even though the entire project occurred during years of solar maximum.

Although the nation is beginning again to send people into space, relatively little attention has been devoted as yet to the radiation belt exposure problem and even less has been addressed to solar flare particle radiation. This subject was recently reviewed by Rust (1982). The ready access of solar particle radiation to the polar regions means that polar-orbiting shuttle flights--and possible future space station flights--will have to be scheduled with an eye on solar activity. Indeed, if a supersonic aircraft were flying a polar route during the August 1972 flare period, it would have had to lower its altitude or all passengers would have received a radiation dose sufficient to produce serious illness (see, for example, the series of reports edited by McKinnon, 1972 and by Vampola, 1979).

Future space activities at geosynchronous altitudes present several forms of radiation hazard to human existence. Even low energy (fraction of an MeV) solar particle radiation has ready access to the entire orbit (Lanzerotti, 1968). Further, temporal changes in the "ambient" trapped environment have occurred in an unpredicted manner, under conditions which are not obviously flare-related. During June 1980 the high energy ( $>5$  MeV) electron fluxes were sufficiently enhanced at geosynchronous orbit that a dosage of  $>1000$  rads would have accumulated by an astronaut in only a few days time (Gaines et al., 1982).

## SPACECRAFT GLOW

One of the more interesting problems concerning environmental influences on technological systems that has recently arisen is the luminescence observed around the space shuttle as it flies at an altitude of  $\sim 200$  km above the Earth (Geophysical Research Letters, February, 1983). This problem has considerable practical importance for scientific and national security considerations. The recent recognition of this luminescence is an important example of the fact that requirements for more sophisticated measurement capabilities reveal that the environment can affect the measurements in unanticipated ways.

At the present writing, the details of the spectra of the glow are poorly characterized. The luminescence around the shuttle is undoubtedly produced by complex chemical and physical processes involved with the impact of the ionized environment on the shuttle surfaces and the atomic physics of the species excited and removed from the surface by the impacting particles. The shuttle has many insulating surfaces (in particular, the famous tiles) and the impacts of charged ions on such insulators (or possibly insulators acting as catalysts with the ions) is not well understood. The atmospheric density can change with solar activity (see Section V), possibly causing changes in the glow measured in a given orbit.

In addition to the induced glows seriously compromising optical measurements, the incident particles can also interact with sensitive surface materials on mirrors and lenses, affecting their optical properties. In extreme cases, for certain coating materials, the incident ions could completely alter, by sputtering or by chemical processes, the surfaces.

## OUTSTANDING QUESTIONS

What are some future problems relating to the effects of solar activity on technological systems? Past experience does not provide much reason for belief that we can anticipate well future problems in this area. Rather the experience has shown us that as the sophistication of technology continues to increase, the sophistication of the knowledge of the natural environment must similarly increase in order to address the largely unanticipated problems. For some present problems, furthermore, many parameters of the environment that are required to answer engineering questions are not now readily available.

For example, in the case of geomagnetic induction effects, the scale sizes of the disturbing fields produced by ionospheric current systems are poorly known. Thus, the magnitude of induction effects in long conductors cannot be easily predicted. Further, it is certainly not possible to predict an expected induction effect from the observation of a specific solar disturbance. As another example, the short- and long-term changes in the radiation belt environments that space spacecraft fly in are not well modeled to better than a factor of two or three. However, a factor of 10% uncertainty is of most interest to spacecraft designers.

## VII. SUMMARY

As we have discussed, solar flares impact the entire solar-terrestrial environment and cause a host of phenomena, some of which have received considerable study in recent years. Yet despite this study, there are some fundamental points, and many missing details of the impact of solar flares, that remain to be understood. We summarize below the important research areas identified in this Section, which should be pursued.

- The flare process yields electromagnetic radiation at all frequencies from radio to gamma rays, energetic particles from solar wind energies to GeV, and enhanced solar wind flow. This radiation and particles provide us with detailed diagnostics of the flare process. Yet with the current observational techniques we do not know the exact location where the flare energy is released, and as a result the exact physical conditions at the flare site. More detailed, higher resolution observations are necessary. These observations can serve as the basis for comprehensive theories of the energy release in solar flares and for shock generation, which do not currently exist, and ultimately for developing models for predicting the occurrence of solar flares.
- Solar flares impact the heliosphere primarily through the enhanced solar wind flow that they generate, and the accompanying shock waves. These flows and shocks can heat the solar wind, modulate and accelerate energetic particles, and cause deviations in the direction of the interplanetary magnetic field, out of the ecliptic plane, which cause geomagnetic disturbances. More comprehensive theories need to be developed to describe the dissipation of enhanced solar wind flows into heat, and the interaction of energetic particles with these flows. For example, theories for galactic cosmic ray interaction are very rudimentary; theories for shock acceleration of energetic particles have developed rapidly in the last several years but need now to be applied in comprehensive models of realistic shocks. Similarly, the mechanisms for causing variations in the interplanetary magnetic field that can drive magnetospheric phenomena need to be understood more fully. Such understanding is an essential component in predicting the impact on the magnetosphere of a given flare.



- Solar flares are responsible for a wide variety of magnetospheric phenomena: they can generate storms and substorms, supply energetic particles to the magnetosphere, and influence the adiabatic and non-adiabatic motion of trapped particles. For some of these phenomena there are conceptual uncertainties that must be resolved; e.g., the microscopic plasma processes that are responsible for solar wind energy conversion to magnetospheric phenomena need to be understood. Other phenomena are conceptually understood, but not in detail, e.g., aspects of storm and substorm dynamics, and details of charged particle transport and energization require study.
- Solar flare EUV and X-rays affect the ionospheric electron content and D-region chemistry; solar flare particles affect D-region ionization, ion/neutral chemistry including stratospheric ozone, and atmospheric electricity. The physics of most of these effects is understood, but their global implications remain to be adequately modeled. For example, an adequate assessment and treatment of global sources, such as magnetospheric electron precipitation, or of global transport of disturbances, needs to be developed. Global modeling is essential if we are ever to determine whether solar flares, or any transient solar effect, influences the weather and/or the climate.
- Solar flares affect many activities in our technological world: radio, satellite and cable communication, power transmission, pipeline corrosion, and manned and unmanned spaceflight. Needed research in this aspect of solar flare impact is difficult to predict, other than to note that as technology becomes more sophisticated our knowledge of the natural environment which can impact that technology must become more detailed. For example, we must improve our modeling of the radiation belt environment of the magnetosphere to satisfy the current and expected needs of spacecraft designers; we must determine the scale-sizes of transient ionospheric and magnetospheric currents to set the design specifications for long, ground-based conductors.

## VIII. REFERENCES

- Akasofu, S. -I., Physics of Magnetospheric Substorms, D. Reidel, Boston, U.S.A., 190, 1977.
- Albertson, V. D., J. M. Thorson, Jr., R. E. Clayton, and S. C. Tripathy, Solar-induced currents in power systems: cause and effects, IEE Trans. Power Apparatus and Systems, vol. PAS-92, p. 471, 1973.
- Anderson, C. W. III, L. J. Lanzerotti, and C. G. MacLennan, Outrage of the L-4 system and the geomagnetic disturbances of August 4, 1972, Bell System Technical Journal, vol. 53, p. 1817, 1974.
- Anderson, K. A., and R. P. Lin, Observation of interplanetary field lines in the magnetotail, J. Geophys. Res., **74**, 3953, 1969.
- Anglin, J. D., W. F. Dietrich, and J. A. Simpson, Conf. Papers 15th Intl. Cosmic Ray Conf. **5**, 43, 1977.
- Arnoldy, R. L., Signature in the interplanetary medium for substorms, J. Geophys. Res., **76**, 5189, 1971

- Arnoldy, R. L., and T. E. Moore, Plasma injection events at synchronous orbit related to positive Dst, J. Geophys. Res., **87**, 77, 1982.
- Aspres, J. D., R. P. Merritt, and S. -I. Akasofu, Effects of geomagnetically induced currents on electric power systems, Northern Engineer, **13**, (3), 34, 1981.
- Axe, G. A., The effects of the earth's magnetism on submarine cables, The Post Office Elect. Engin. J., **61**, Part 1, 37, 1968.
- Baker, D. N., P. R. Higbie, R. D. Belian, and E. W. Hones, Jr., Do Jovian electrons influence the terrestrial outer radiation zone? Geophys. Res. Lett., **6**, 531, 1979.
- Baker, D. N., R. D. Zwickl, S. J. Bame, E. W. Hones, Jr., B. T. Tsurutani, E. J. Smith, and S. -I. Akasofu, An ISEE high time resolution study of interplanetary parameter correlations with magnetospheric activity, to appear in J. Geophys. Res., 1983.
- Barouch, E., and L. F. Burlaga, Causes of Forbush decreases and other cosmic ray variations, J. Geophys. Res., **80**, 449, 1975.
- Barlow, W. H., On the spontaneous electrical currents observed in the wires of the electric telegraph, Phi. Trans. R. Soc., **61**, 1849.
- Brooks, J., The subtle storm, New Yorker Magazine, p. 39, Feb. 19, 1959.
- Campbell, W. H., and J. E. Zimmerman, Induced electric currents in the Alaska oil pipeline measured by gradient fluxgate and SQUID magnetometer, IEEE Trans. Geosci and Remote Sensing, **GE-18**, 244, 1980.
- Castelli, J. P., and D. A. Guidice, The radial event associated with the polar cap absorption event of 2 November 1969, Proc. COSPAR Sym. Solar Particle Events of November 1969, ed. by J. C. Ulrick, AFCRL Report 72-0474, 1972.
- Chapman, S., Outline of a theory of magnetic storms, Proc. Roy. Soc., (London), **A95**, 61, 1918.
- Chen, G., and T. P. Armstrong, Acceleration of charged particles in oblique MHD shocks, Proc. Int. Cof. Cosmic Rays 14th, **5**, 1814, 1975.
- Chiu, Y. T., and J. M. Straus, Rayleigh-Taylor and wind-driven instabilities of the nighttime equatorial ionosphere, J. Geophys. Res., **84**, 3283, 1979.
- Chupp, E. L., High energy particle acceleration in solar flares, observational evidence, Solar Phys., in press, 1983.
- Chupp, E. L., D. J. Forrest, J. M. Ryan, J. Heslin, C. Peppin, K. Pinkau, G. Kanbach, E. Reiger, and G. H. Share, A direct observation of solar neutrons following the 0118 U.T. flare on June 21, 1980, Astrophys. J., **263**, L95, 1982.
- Collard, H. R., J. D. Mihalov, and J. H. Wolfe, Radial variation of the solar wind speed between 1 and 15 AU, J. Geophys. Res., **87**, 2203, 1982.

- Cook W. R., E. C. Stone, and R. E. Vogt, Elemental composition of solar energetic nuclei, Astrophys. J. (Letters), **238**, L97, 1980.
- Davidson, W. F., The magnetic storm of March 24, 1940 -- effects in the power system, Edison Electric Institute Bulletin 365, July, 1940.
- DeForest, S. Electrostatic charging on spacecraft in Solar System Plasma Physics, ed. L. J. Lanzerotti, C. F. Kennel, and E. N. Parker, North-Holland Pub. Co., vol. 3, p. 336, 1979.
- Donnelly, R. F., Empirical models of solar flare X-ray and EUV emissions for use in studying their E and F region effects, J. Geophys. Res., **81**, 4745-4753, 1976.
- Donnelly, R. F., K. Davies, R. N. Grubb, and R. B. Fritz, F-region enhancements induced by solar flares, The Geophysical Use of Satellite Beacon Observations, ed. M. Mendillo, 345-359, 1976.
- Duijveman, A., P. Hoyng, and M. E. Machado, X-ray imaging of three solar flares during the impulsive phase, Solar Phys., **81**, 137, 1982.
- Emmons, S., Sunspots playing tricks with radios, Los Angeles Times, p. CC-1, February 13, 1979.
- Fairfield, D. H., and L. J. Cahill, Jr., Transition region magnetic field and polar magnetic disturbances, J. Geophys. Res., **71**, 155, 1966.
- Fan, C. Y., G. Gloeckler, and D. Hovestadt, The composition of heavy ions in solar energetic particle events, Univ. of Md., preprint, 1983.
- Feldman, W. C., J. R. Asbridge, S. J. Bame, J. T. Gosling, and R. D. Zwickl, Electron heating at interplanetary shocks, Solar Wind V., ed. by M. M. Neugebauer, NASA, Washington, DC, 1983.
- Fisk, L. A.,  $^3\text{He}$ -rich flares: a possible explanation, Astrophys. J., **224**, 1048, 1978.
- Fisk, L. A., and M. A. Lee, Shock acceleration of energetic particles in corotating interaction regions in the solar wind, Astrophys. J., **237**, 620, 1980.
- Forman, M. A., R. Ramaty, and E. G. Zweibel, The acceleration and propagation of solar flare energetic particles, Chap. 1 The Physics of the Sun, ed. T. E. Holzer, D. Mihalas, P. A. Sturrock, and R. K. Ulrich, 1982.
- Gaines, E. E., R. W. Nightingale, W. L. Imhof, and J. B. Reagan, Enhanced radiation doses to high-altitude spacecraft during June 1980, ISEE Trans. Nucl Sci., **NS-28**, No. 6, 1981.
- Garrard, T. L., E. C. Stone, and R. E. Vogt, The isotopes of H and He in solar cosmic rays, High Energy Phenomena on the Sun (NASA Sp-342), p. 341, 1973.
- Garrett, H. B., A. J. Dessler, and T. W. Hill, Influence of solar wind variability on geomagnetic activity, J. Geophys. Res., **79**, 4603, 1974.
- Gary, S. P., and T. E. Eastman, The lower hybrid drift instability at the magnetopause, J. Geophys. Res., **84**, 7378, 1979.

- Gazis, P. R., and A. J. Lazarus, The radial evolution of the solar wind, 1-10 AU, Solar Wind V, ed. by M. Neugebauer, NASA, Washington, DC, 1983.
- Gendrin, R., Diffusion processes in the magnetopause boundary layer, Geophys. Res. Lett., 1983.
- Germaine, L. W., Magnetic storm of March 24, 1940 -- effects in communications systems, Edison Electric Institute Bulletin, May 7, 1940.
- Gloeckler, G., H. Weiss, D. Hovestadt, F. M. Ipavich, B. Klecker, L. A. Fisk, M. Scholer, C. Y. Fan, and J. J. O'Gallagher, Observations of the ionization states of energetic particles accelerated in solar flares, Conf. Papers 17th Intl. Cosmic Ray Conf., 3, 136, 1981.
- Gold, T., Magnetic storms, Space Sci. Rev., 1, 100, 1962.
- Goldstein, B. E., and J. R. Jokipii, Effects of stream-associated fluctuations upon the radial variations of average solar wind parameters, J. Geophys. Res., 82, 1095, 1977.
- Heikkila, W. J., Impulsive plasma transport through the magnetopause, Geophys. Res. Lett., 9, 159, 1982.
- Hoyng, P. A., Duijveman, M. E. Machado, D. M. Rust, Z. Svestka, A. Boclee, C. deJager, K. J. Frost, H. Lafleur, G. M. Simnett, H. F. vanBeek, and B. E. Woodgate, Origin and location of the hard X-ray emission in a two-ribben flare, Astrophys. J., 246, L155, 1981.
- Hudson, P. D., Reflection of charged particles by plasma shocks, Mon. Not. R. Astron. Soc., 131, 23, 1965.
- Hundhausen, A. J., Coronal Expansion and Solar Wind, Springer, Heidelberg, 1972.
- Kelley, M. C., G. Haerendel, H. Kappler, A. Valenzuela, B. B. Balsley, P. A. Carter, W. Eckland, C. W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread-F, Geophys. Res. Lett., 3, 448, 1976.
- Kennel, C. F., F. L. Scarf, F. V. Coroniti, E. J. Smith, and D. A. Gurnett, Non-local plasma turbulence associated with interplanetary shocks, J. Geophys. Res., 87, 17, 1982.
- King, J. J., Solar proton fluences for 1977-83 space missions, J. Spacecr. Rockets, 11, 401, 1974.
- Kolasinki, W. A., J. B. Blake, J. K. Anthony, W. E. Price, and E. C. Smith, Simulation of cosmic-ray induced soft errors and latchup in integrated-circuit computer memories, Proc. IEEE Trans. Nucl. Sci., 1979.
- Krimigis, S. M., E. T. Sarris, and T. P. Armstrong, Observation of Jovian electron events in the vicinity of earth, Geophys. Res. Lett., 2, 561-564, 1975.
- Lanzerotti, L. J., Penetration of solar protons and alphas to geomagnetic equator, Phys. Rev. Lett., 21, 929, 1968.
- Lee, M. A., A coupled hydromagnetic wave excitation and in acceleration upstream of the earth's bow shock, J. Geophys. Res., (in press), 1983.

- Lee, M. A., and L. A. Fisk, Shock acceleration in the heliosphere, Space Sci. Rev., **32**, 205, 1982.
- Lemaire, J., Impulsive penetration of filamentary plasma elements into the magnetospheres of the Earth and Jupiter, Planet. Space Sci., **26**, 887, 1977.
- Lin, R. P., Non-relativistic solar electrons, Space Sci. Rev., **16**, 189, 1974.
- Lin, R. P., and H. S. Hudson, Non-thermal processes in large solar flares, Solar Phys., **50**, 153, 1976.
- Lin, R. P., R. A. Mewaldt, and M. A. I. Van Hollebeke, The spectrum of 20keV-20MeV electrons accelerated in large solar flares, Astrophys. J., **253**, 949, 1982.
- Mariska, J. T., and E. S. Oran, The E and F region ionospheric response to solar flares, J. Geophys. Res., **86**, 5868-5872, 1981.
- Mason, G. M., L. A. Fisk, D. Hovestadt, and G. Gloeckler, A survey of  $\sim 1$  MeV nucleon<sup>-1</sup> solar flare particle abundances  $1 \leq Z \leq 26$ , during the 1973-1977 solar minimum period, Astrophys. J., **239**, 1070, 1980.
- Mason, G. M., G. Gloeckler, and D. Hovestadt, Temporal variations of nucleonic abundances in solar flare energetic particle events I: well-connected events, Univ. of Md. preprint, 1983.
- Ma Sung, L. S., G. Gloeckler, and D. Hovestadt, Measurement of ionization states for low energy heavy ions in solar flare particle events, Conf. Papers, 17th Int. Cosmic Ray Conf., **10**, 45, 1981a.
- Ma Sung, L. S., G. Gloeckler, and D. Hovestadt, Ionization of heavy elements observed in the 1974 May 14-15 anomalous solar event, Astrophys. J. (Letters), **245**, L45, 1981b.
- McDonald, F. B., J. H. Trainor, and W. R. Webber, Pioneer and Voyager observations of Forbush decreases between  $\sim 6$  and 24 AU, 17th Int. Conf. on Cosmic Rays, **11**, 147, 1981.
- McGuire, R. E., T. T. von Rosenvinge, and F. B. McDonald, A survey of solar proton and alpha differential spectra between 1 and  $>400$  MeV/nucleon, Conf. Papers 17th Intl. Cosmic Ray Conf., **3**, 65, 1981.
- McKinnon, J. A., August 1972 Solar activity and related geophysical effects, NOAA TMERL SEL-22, USDC, Boulder, Co., 1972.
- Oran, E. S., and T. R. Young, Numerical modeling of ionospheric chemistry and transport processes, J. Phys. Chem., **81**, 2463-2467, 1977.
- Oran, E. S., L. Baker, and T. R. Young, The Effects of Solar Flares on the Ionosphere, Memo. 3296, Naval Research Laboratory, Washington, DC, 20375, 1976.
- Parker, E. N., Dynamics of the interplanetary gas and magnetic fields, Astrophys. J., **128**, 664, 1958.

- Parker, E. N., Sudden expansion of the corona following a large solar flare and the attendant magnetic field and cosmic ray effects, Astrophys. J., **133**, 1014, 1961.
- Paulikas, G. A., The space radiation environment: effects on space systems in Solar System Plasma Physics, ed. L. J. Lanzerotti, C. F. Kennel, and E. N. Parker, North-Holland Pub. Co., vol. 3, p. 330, 1969.
- Paulikas, G. A., and J. B. Blake, Penetration of solar protons to synchronous altitude, J. Geophys. Res., **74**, 2161-2168, 1969.
- Paulikas, G. A., and L. J. Lanzerotti, Impacts of geospace effects on terrestrial technology, Astronautics and Aeronautics, **20**, 42, July/August 1982.
- Pesses, M. E., and R. B. Decker, On the acceleration of ions by interplanetary shock waves II. Multiple encounter considerations, submitted to J. Geophys. Res., 1983.
- Potter, D. W., Acceleration of the electrons by interplanetary shock, J. Geophys. Res., **88**, 11111, 1981.
- Prescott, G. B., The great auroral exhibition of August 28th to September 4th, 1859; 1. Observations made at Boston, Mass., and its vicinity: Am. J. Sci. Arts, **29**, 92, 1860.
- Ramaty, R., Nuclear processes in solar flares, Chap 3, The Physics of the Sun, ed. T. E. Holzer, D. Mihalas, P. A. Sturrock, and R. K. Ulrich, 1982.
- Ramaty, R., S. A. Colgate, G. A. Dulk, P. Hoyng, J. W. Knight, R. P. Lin, D. B. Melrosa, C. Paizes, F. Orrall, P. R. Shapiro, D. F. Smith, and M. I. A. Van Hollebeke, Energetic particles in solar flares, in Proc. of Second Skylab Workshop on Solar Flares, Chap. 4, Boulder, Colorado, 1979.
- Reagan, J. B., Ionization processes, in Dynamical and Chemical Coupling Between the Neutral and Ionized Atmosphere, edited by B. Grandal and J. A. Holtet, 145, D. Reidel, Hingham, Mass., 1977.
- Reagan, J. B., and T. M. Watt, Simultaneous satellite and radar studies of the D-region ionosphere during the intense solar particle events of August 1972, J. Geophys. Res., **81**, 4579, 1976.
- Reagan, J. B., R. E. Meyerott, R. W. Nightingale, R. C. Gunton, R. G. Johnson, J. E. Evans, and W. L. Imhof, Effects of the August 1972 solar particle events on stratospheric ozone, J. Geophys. Res., **86**, 1473, 1981.
- Reagan, J. B., R. E. Meyerott, J. E. Evans, W. L. Imhof, and R. G. Joiner, The effects of energetic particle precipitation on the atmospheric electric circuit, J. Geophys. Res., **88**, No. C6, 3869, 1983.
- Reinhard, R., P. Van Nes, T. R. Sanderson, K-P. Wenzel, E. J. Smith, and B. T. Tsurutani, A statistical study of interplanetary shock associated proton intensity increases, 18th Int. Cosmic Ray Conf., 1983.

- Rosenberg, T. J., P. B. Morris, and L. J. Lanzerotti, Excitation of magnetospheric hydromagnetic waves by solar-flare-induced changes in ionospheric conductivity, Phys. Rev. Letters, **47**, 1343, 1981.
- Russell, C. T., E. J. Smith, B. T. Tsurutani, J. T. Gosling, and S. J. Bame, Multiple spacecraft observations of interplanetary shocks: characteristics of the upstream ULF turbulence, Solar Wind V, ed. by M. M. Neugebauer, NASA, Washington, DC, 1983.
- Rust, D. M., Solar flares proton showers and the Space Shuttle, Science, **216**, 939, 1982.
- Scannapieco, A. J., and S. L. Ossakow, Nonlinear equatorial spread-F, Geophys. Res. Lett., **3**, 451, 1976.
- Schildge, J. P., and G. L. Siscoe, A correlation of occurrence of simultaneous sudden magnetospheric compressions and geomagnetic bay onsets with selected geophysical indices, J. Atmos. Terr. Phys., **32**, 1819, 1970.
- Scholer, M., F. M. Ipavich, G. Gloeckler, D. Hovestadt, Acceleration of low energy protons and alpha particles at interplanetary shock waves, J. Geophys. Res., **88**, 1977, 1983.
- Schulz, M., and L. J. Lanzerotti, Particle Diffusion in the Radiation Belts, Springer, Heidelberg, 1974.
- Sciambi, R. K., G. Gloeckler, C. Y. Fan, and D. Hovestadt, Emission of nearly stripped carbon and oxygen from the sun, Astrophys. J., **214**, 316, 1977.
- Siscoe, G. L., V. Formisano, and A. J. Lazarus, Relation between geomagnetic sudden impulses and solar wind pressure changes, an experimental investigation, J. Geophys. Res., **73**, 4869, 1968.
- Slothower, J. C., and V. D. Albertson, The effects of solar magnetic activity on electric power systems, J. Minn. Acad. Sci., vol. **34**, p. 94, 1967.
- Smith, D. F., First phase acceleration in solar flares, in Particle Acceleration Mechanisms in Astrophysics, Am. Inst. Phys., **155**, 1979.
- Smith, E. J., Observed properties of interplanetary shocks, EOS, **63**, 398, 1982.
- Smith, E. J., and A. Barnes, Spatial dependencies in the distant solar wind: Pioneers 10 and 11, Solar Wind V, ed. by M. M. Neugebauer, NASA, Washington, DC 1983.
- Swider, W., T. J. Keneshea, and C. I. Foley, An SPE-disturbed D-region model, Planet. Space Sci., **26**, 883, 1978.
- Syrovatskii, S. E., Pinch sheets and reconnection in astrophysics, Annual Revs. Astron. Astrophys., **19**, 163, 1981.
- Taur, R. R., Ionospheric scintillations at 4 and 6 GHz, COMSAT Technical Review, **3**, 145, 1973.
- Taur, R. R., Gigahertz ionospheric scintillation, COMSAT Technical Review, **4**, 1974.

- Taur, R. R., Simultaneous 1.5 and 4 GHz ionospheric scintillation, Radio Sciencies, 1976.
- Taur, R. R., Scintillation of communication satellite signals at GHz frequencies, in Solar System Plasma Physics, ed. L. J. Lanzerotti, C. F. Kennel, and E. N. Parker, North-Holland Pub Co., Vol. 3, p. 320, 1979.
- Tsurutani, B. T., and C. I. Meng, Interplanetary magnetic field variations and substorm activity, J. Geophys. Res., **77**, 2964, 1972.
- Tsurutani, B. T., and R. M. Thorne, Diffusion processes in the magnetopause boundary layer, Geophys. Res. Lett., 1247, 1982.
- Tsurutani, B. T., E. J. Smith, and D. E. Jones, Waves observed upstream of interplanetary shocks, J. Geophys. Res., 1983a.
- Tsurutani, B. T., C. T. Russell, J. J. King, R. D. Zwickl and R. P. Lin, A kinky, inclined, heliospheric current sheet: magnetospheric substorms, submitted to Geophys. Res. Lett., 1983b.
- Vampola, A. L., User requirements of solar terrestrial predictions for spacecraft applications, Solar-Terrestrial Predictions Proc., ed. R. F. Donnelly, NOAA-SEL USDC Boulder Co., 2, 1, 1979.
- Varvoglis, H., and K. Papadopoulos, Selective non-resonance acceleration of  $^3\text{He}^{++}$  and heavy ions by  $\text{H}^+$  cyclotron waves, Astrophys. J., (Letters), in press, 1983.
- Vasyliunas, V. M., Theoretical models of magnetic field line merging 1, Rev. Geophys. Sp. Phys., **13**, 303, 1975.
- Wild, J. P., S. F. Smerd, and A. A. Weiss, Solar bursts, Ann. Rev. Astron. Astrophys., **1**, 291, 1963.
- Winckler, J. R., L. Peterson, R. Hoffman, and R. Arnoldy, Auroral X-rays, cosmic rays, and related phenomena during the storm of February 10, 1958, J. Geophys. Res., **64**, 597, 1959.



# CHAPTER 10

## COLLISIONLESS SHOCK WAVES IN THE SOLAR TERRESTRIAL ENVIRONMENT

### WORKING GROUP MEMBERS

E. Greenstadt, Chairman  
*TRW*

V. Formisano  
*I.F.S.I./CNR*

C. Goodrich  
*University of Maryland*

J. T. Gosling  
*Los Alamos National Laboratory*

M. Lee  
*University of New Hampshire*

M. Leroy  
*DESPA Observatoire de Meudon*

M. Mellott  
*University of California*

K. Quest  
*Los Alamos National Laboratory*

A. E. Robson  
*Naval Research Laboratory*

P. Rodriguez  
*Naval Research Laboratory*

J. Scudder  
*NASA/Goddard Space Flight Center*

J. Slavin  
*Jet Propulsion Laboratory*

M. Thomsen  
*Los Alamos National Laboratory*

D. Winske  
*Los Alamos National Laboratory*

C. S. Wu  
*University of Maryland*

**CHAPTER 10****COLLISIONLESS SHOCK WAVES IN THE SOLAR TERRESTRIAL ENVIRONMENT**

I.	Introduction	10-3
II.	Natural Shocks	10-4
III.	Quasi-Perpendicular Supercritical Processes	10-16
IV.	Quasi-Perpendicular Subcritical Processes	10-30
V.	Quasi-Parallel Processes	10-31
VI.	Ion Acceleration	10-35
VII.	Outstanding Problems	10-43
VIII.	Investigative Avenues	10-44
IX.	References	10-50

## I. INTRODUCTION

Shock waves are created by the nonlinear steepening of compressive wavemodes in a fluid. Typically such steepening occurs when a disturbance travels through the fluid at a speed higher than the characteristic speed with which small amplitude compressive waves can propagate. The shock formed in front of such a disturbance is the means by which the disturbance communicates with the ambient plasma into which it is propagating. At the shock the flow normal to the shock must be changed from super-"sonic" upstream to sub-"sonic" downstream, where "sonic" refers to the characteristic speed of the small amplitude pressure signals. This plasma speed is usually the magnetoacoustic speed. The loss of streaming energy represented by this slowing of the flow at the shock is converted into other forms of energy. Broadly speaking in fact, a shock may be defined as the entire region over which any portion of the conversion takes place. In space plasmas this conversion must be accomplished via collisionless dissipation mechanisms. Identifying and understanding dissipation mechanisms is therefore a central concern of many shock studies.

Laboratory experiments, numerical simulations, and space observations all indicate that the modes of dissipation at a shock depend sensitively upon upstream flow conditions, particularly the Mach number of the flow,  $M$ , and the angle between the local shock normal and the upstream magnetic field,  $\theta_{Bn}$ , and the ratio of thermal to magnetic internal energy,  $\beta$ . Shocks with Mach numbers below some critical value,  $M_c$ , are known as subcritical shocks, whereas those with  $M > M_c$  are known as supercritical.  $M_c$  itself depends upon such things as the upstream plasma  $\beta$  and  $\theta_{Bn}$ , but is usually within the range 2-3. The fundamental difference between these two classes of shocks is that dispersion and anomalous resistivity provides most of the dissipation for subcritical shocks whereas ion reflection at the shock plays an increasingly important role in the dissipation process as the Mach number increases above  $M_c$ . A fundamental change in shock structure occurs also at  $\theta_{Bn} \sim 45^\circ$  which is most evident as an increase in turbulence and the lack of a well defined shock structure when  $\theta_{Bn}$  is less than  $45^\circ$ . Thus,  $\theta_{Bn}$  provides a rough dividing line between two further classes of shocks: quasi-perpendicular ( $\theta_{Bn} \gtrsim 45^\circ$ ) and quasi-parallel ( $\theta_{Bn} \lesssim 45^\circ$ ).

Small amplitude waves can propagate through a plasma with different characteristic speeds depending upon the mode of propagation. This difference leads to a further classification of shock structure - slow, intermediate, and fast which, correspond to disturbances exceeding respectively the slow magnetoacoustic speed, the Alfvén speed, and the fast magnetoacoustic speed. Only fast shocks have been studied extensively; however, slow shocks have been observed in the solar wind and may play an important role in field line reconnection processes. In this report we concentrate exclusively on fast shocks, unless stated otherwise.

Collisionless shocks commonly contain a structural element known as the foreshock which is not present in ordinary collisional shocks. The foreshock extends for a considerable distance upstream from the main shock transition and contains suprathermal ions and electrons reflected at the shock or transmitted through it from the downstream region. These ions and electrons excite a variety of hydromagnetic and plasma waves, which in turn scatter the particles. As the wave speeds are generally lower than the convection speed of the upstream plasma flow, the waves are subsequently convected back toward the shock. Ultimately, scattering of the initial "seed" populations of reflected and/or leaked particles off of these waves both up and downstream from the shock can produce a significant acceleration of a small fraction of the ions to very high energy. Such acceleration appears to be a ubiquitous feature of all shock waves observed within the heliosphere.

Progress in our knowledge of collisionless shocks, with particular attention to the important contributions made by satellite instruments, has been reviewed regularly (Dobrowolny and Formisano, 1973; Greenstadt and Fredricks, 1979; Formisano, 1981; Russell and Greenstadt, 1983). We do not intend a historical or general review here, but rather a summary of selected information pertaining to the current status of our understanding of the physical processes in shocks. Much has been learned lately; even more is being studied and prepared for publication as this is written. Exposing the newest results and their relationships to each other is a major aim of this report. We focus on the large and small scale physics of stable structures, ignoring the transient formation or development of shocks.

The knowledgeable reader will find an unorthodox order of presentation here: Tradition dictates that shock descriptions begin with simple (cold, laminar) cases and proceed to the more complicated (hot, turbulent) examples. So much is being learned now about supercritical shocks, which are, after all, the most common ones offered for investigation in the solar wind, that the Working Group found itself more comfortable explaining these relatively complex, but familiar, structures first and then approaching subcritical shocks in terms of the elimination of factors unimportant to the physics of the latter.

In the following sections we outline the context, especially the natural one, in which critical questions associated with collisionless shock structure and maintenance arise, our present understanding of the essential processes responsible for shock phenomenology, and a list of unanswered questions which still frustrate our understanding of these fascinating and ubiquitous structures.

## II. NATURAL SHOCKS

### OCCURRENCE

Plasma shock waves, not necessarily all collisionless, occur throughout the solar system and apparently in stellar and galactic settings as well. Plasma in the solar-terrestrial environment is collisionless outside the Sun's atmosphere, where shock waves are produced by high speed solar mass ejections, by the steepening of high speed solar wind streams, by the flow of the solar wind past diamagnetic planetary envelopes and planetary cometary halos, by the motion of the heliosphere itself through the interstellar medium, and perhaps by the rapid release of energy via reconnection wherever it takes place inside, outside, or at the boundaries of magnetospheres. We infer from telescopic observations that shock waves must be produced by events of explosive force all over the universe, and we may reasonably expect that the conditions giving rise to shocks within the solar-terrestrial environment are reproduced inside other stellar systems. Indeed, we believe that some of the information we receive at Earth about the distant cosmos, e.g., through cosmic rays, is generated by particle acceleration even in relatively weak shocks so that there is a direct connection between what we learn directly in our own vicinity and what we can learn about the larger astrophysical scene.

It may be argued, in addition, that shocks arising in plasmas of human contrivance are as natural as any, and, in fact, man-made conditions resulting in shocks are not altogether unusual. They are routine in nuclear explosions and arise in the laboratory, intentionally or unintentionally, in connection with experiments on thermonuclear plasmas. Certainly, however, the most common occurrence of shocks, and the one most accessible to repeated observation and detailed measurement with instruments small compared to shock-structural scales, is at our doorstep to the universe, in the nearby solar wind.

## GLOBAL FEATURES

### MHD Configurations

Here, we assess the success of theory in describing the global aspects of these shock waves with an emphasis on bowshocks as the observational paradigm. Interplanetary shocks have also been the subject of extensive experimental and theoretical studies (e.g., Shea et al., 1982), but each event is unique and the number of observations for a given event cannot compare with the bowshock database.

The riches of the bowshock database, and its deficiencies, may be appreciated by examining Figure 10-1. The shock is diagrammed in three dimensions as three curves representing the intersections of a hyperboloid of revolution with the  $x$ - $y$ ,  $x$ - $z$ , and  $y$ - $z$  planes. The hyperboloid is an observed average surface (Greenstadt, 1972, Appendix), located and scaled according to average solar wind parameters; as noted in the figure, the average shock is divided roughly at its vertical,  $x$ - $z$  centercurve (actually at the dashed curve, a little east of the centercurve), into quasi-perpendicular and quasi-parallel structures. The division is determined by the direction of the average interplanetary magnetic field (IMF). Each dot marks the point on the average surface where a radial line from the origin to an observed shock encounter crossed this surface; i.e., each dot shows the polar coordinates of a shock crossing on the average surface, but not the correct radial distance of the crossing, which depended on the prevailing solar wind conditions at the time the crossing was recorded. The figure provides a survey of the geometric coverage of the Earth's bowshock and give a rough impression of likely relative numbers of quasi-perpendicular and quasi-parallel observations. The figure does not give any impression of the range of shock expansions or contractions over which crossings really occurred.

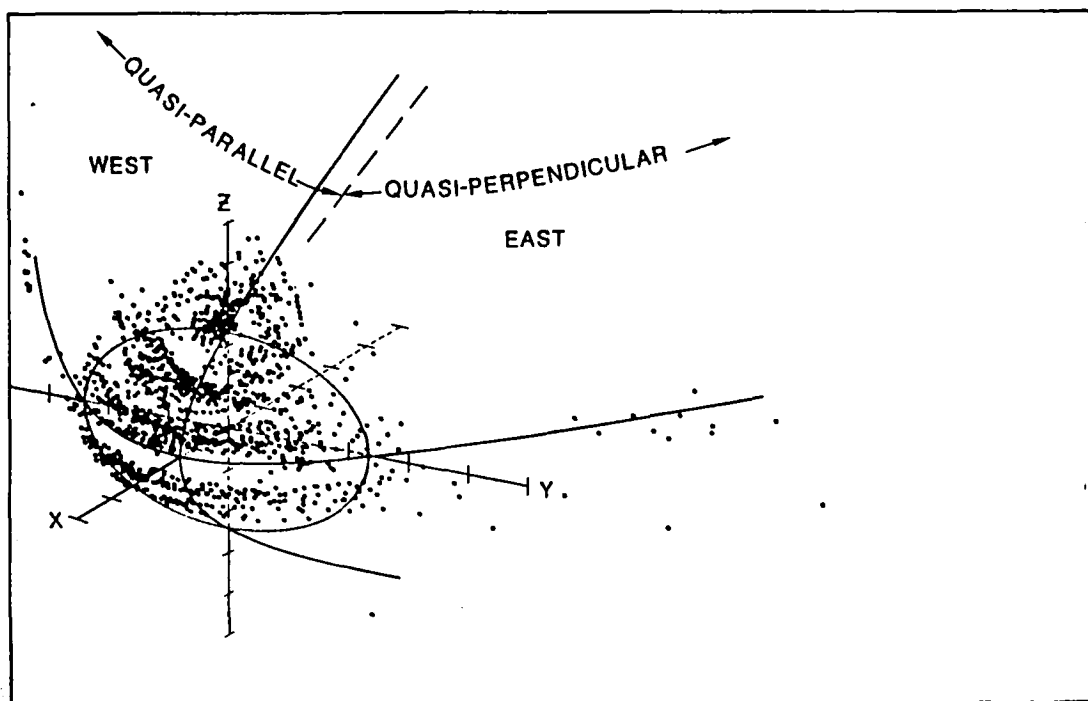


Figure 10-1. Angular locations on a mean shock surface of shock encounters by IMPs 1, 2, 6, & 8; OGOs 1 & 5; EXPs 33 & 35; HEOS 1 & 2; PROGNOZ 1, 2, 4, 5, & 7; and ISEE 1, compiled by Slavin (1983).

Clearly, the subsolar shock ( $x > 0$ ) has been sampled very well in many directions, especially in the northern ( $z > 0$ ) low, mid, and polar latitudes. Just as clearly, the shock has not been sampled uniformly; its flanks and southern subecliptic and polar areas ( $z < 0$ ) are relatively deficient in coverage. The experienced analyst will recognize, given the variability of the solar wind, that, on one hand, a large array of samples like that in Figure 10-1 is necessary to obtain well populated subsets constant in any parameter, but, on the other hand, study of global characteristics free of transient effects and applicable to the class of planetary shocks requires normalization of every observation to a prescribed state. Granting the appropriate use of normalizing techniques, the subsolar region is visibly overdetermined, and has been well modeled. The deficiencies of the sample mainly degrade the accuracy of shock models along the far flanks and diminish confidence in their north-south symmetry.

Figure 10-2 presents a schematic view of solar wind flow about the terrestrial magnetosphere as we know it today. The high speed of the solar wind and the largely impenetrable nature of the magnetosphere to the solar wind create a strong bowshock to divert the flow. As will be discussed in a later section, the angle between the interplanetary magnetic field and the shock normal exerts great influence over the shock structure and controls the upstream regions to which reflected particles have access.

Within the magnetosheath there is considerable convected turbulence (Fairfield, 1976) as well as locally generated waves (Tsurutani et al., 1982a). Adjacent to the dayside magnetopause there is a thin,  $0.2-0.4 R_E$ , plasma depletion layer (Crooker et al., 1979). The forward magnetopause is also the site of the interaction between the geomagnetic field that creates and maintains the Earth's magnetosphere and the magnetosheath plasma. The interaction mechanisms, such as reconnection, are discussed in other portions of this report.

Although not so well sampled as the terrestrial bowshock and magnetosheath, studies at other planets offer an opportunity to examine Mach numbers and IMF orientations that are far less common at 1 AU. Figure 10-3 displays bowshock location models for Venus, Earth, Mars, Jupiter, and Saturn (Slavin and Holzer, 1981; Slavin et al., 1983b) plotted in terms of shock stand-off distance. The planets have also been displayed to provide a measure of the absolute scaling. The Parker spiral orientation of the interplanetary magnetic field in the ecliptic is shown in the upper left hand corner. While the solar wind Mach numbers increase somewhat with growing distance from the Sun and the average IMF direction becomes increasingly azimuthal (e.g., Slavin and Holzer, 1981), the differences in bowshock shape and stand-off distance in Figure 10-3 have all been attributed largely to the nature of the interaction between each planet and the solar wind plasma (e.g., magnetospheric shape, charge exchange or photoionization involving neutrals in the magnetosheath). The alternations in the regions of quasi-perpendicular/quasi-parallel shock structure are thought, but not yet proven, to be secondary in terms of the overall flow field. The success of solar planetary interaction modeling will ultimately require a mature understanding of the shocked solar wind and hence shock structure as a function of all the different upstream conditions that can occur in the solar system.

Given the figure of the magnetosphere and suitable magnetopause boundary conditions, theoretical modeling of bowshock position and strength is ideally a time-dependent multi-fluid MHD problem with equations of state and jump conditions for each species dictated by the processes responsible for thermalization (Spreiter, 1976). Additional terms describing the energy and momentum carried by reflected particles and downstream turbulence may also be necessary (Chao and Wiskerchen, 1974). Unfortunately, the computational codes that have been applied for this situation do not closely approach the ideal. All are single fluid despite the very different ion and electron temperature jump conditions that have been measured across collisionless shocks in space (Montgomery et al., 1970; Bame et al., 1979). The only code which self-consistently

## SOLAR WIND FLOW ABOUT THE TERRESTRIAL MAGNETOSPHERE

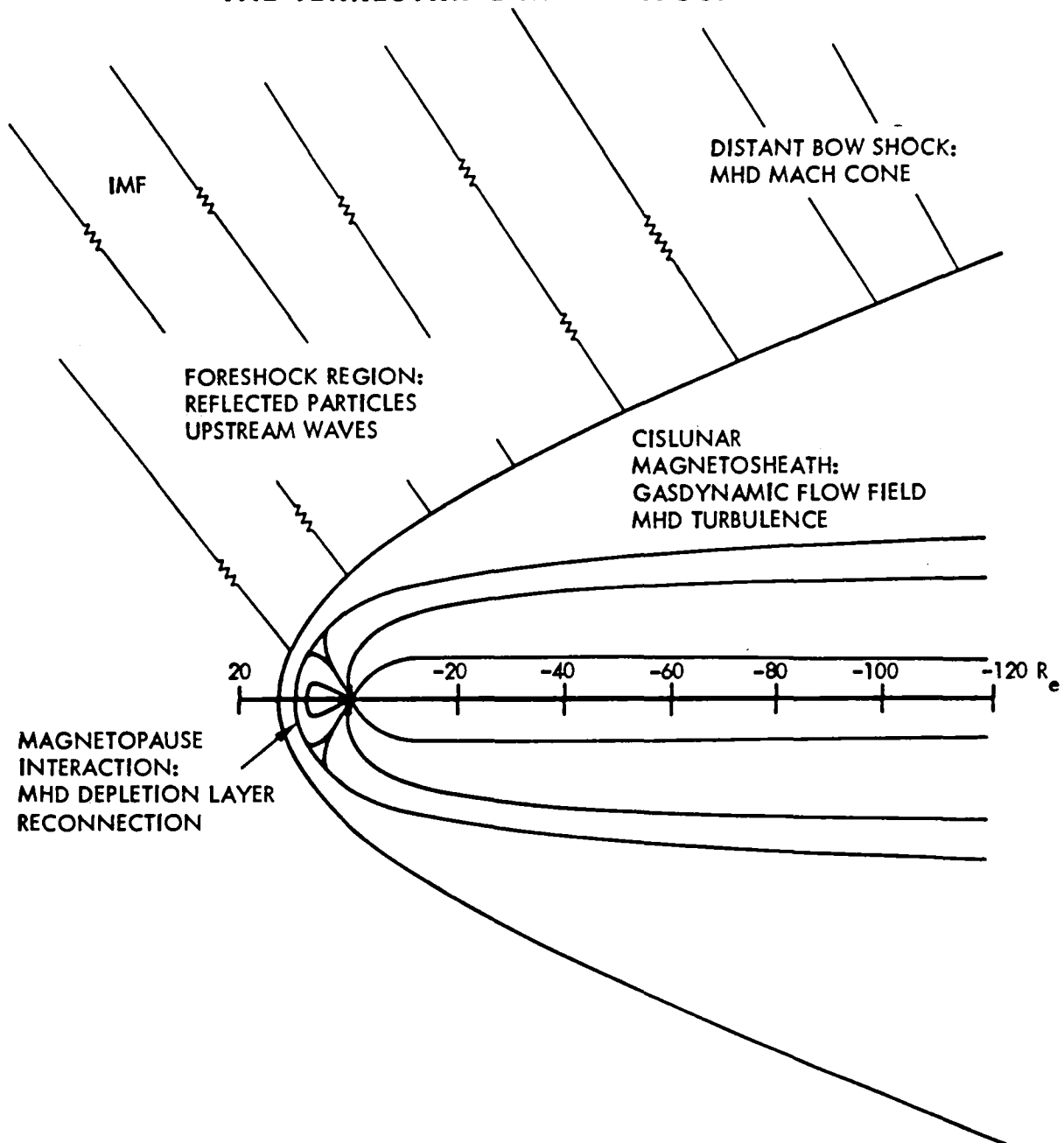


Figure 10-2. Major regions of the solar wind's interaction with the Earth's magnetosphere. As indicated in the text, a combination of gasdynamic, magnetohydrodynamic, and plasma kinetic theory is used in modeling the flow of solar wind about the magnetosphere.

## INTERPLANETARY MAGNETIC FIELD-BOW SHOCK GEOMETRIES

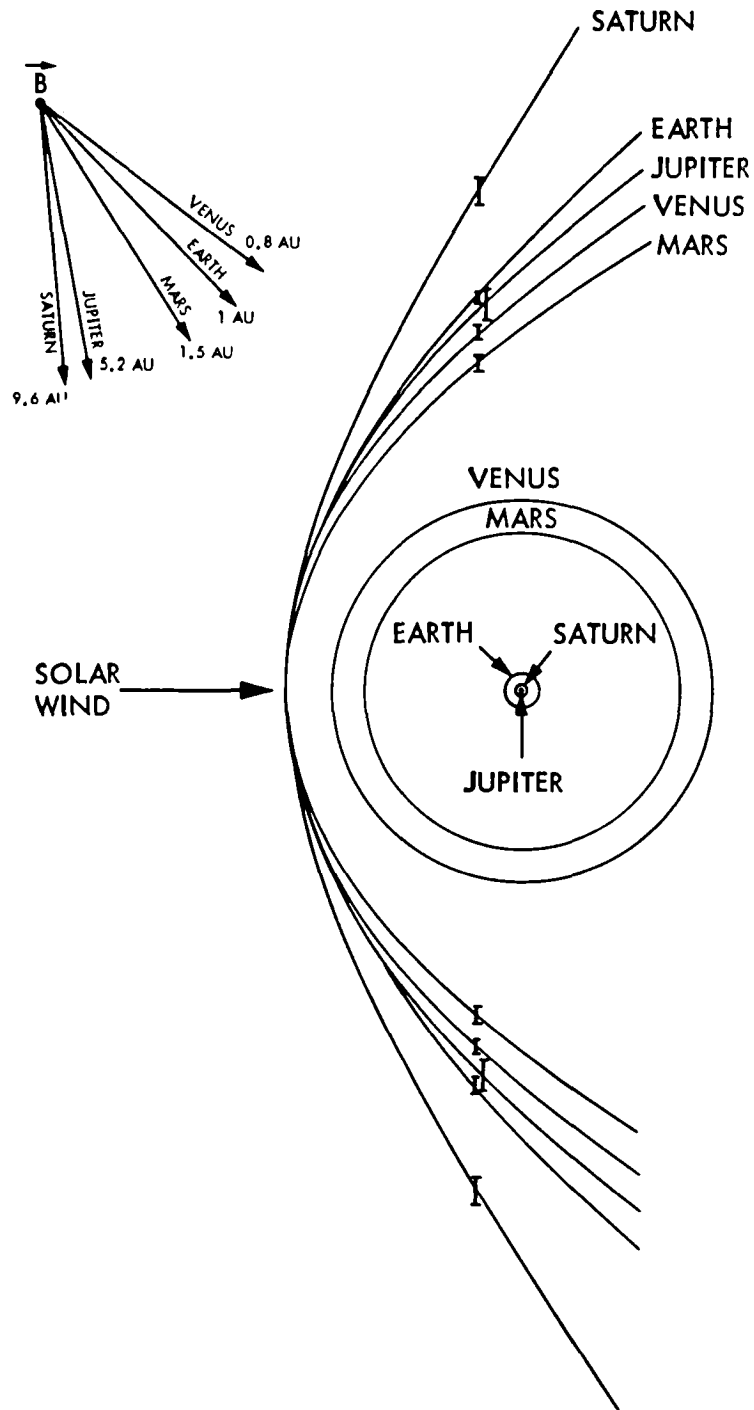


Figure 10-3. Average bowshock surfaces at Venus, Earth, Mars, Jupiter, and Saturn in units which set the shock stand-off distance equal to unity. Average interplanetary magnetic field orientation at the orbit of each planet is shown at upper left.



includes the magnetic field and has been run for a variety of upstream Mach numbers is the field aligned flow model of Spreiter and Rizzi (1974). Recently, MHD simulation codes have been developed for studying the solar wind interaction with the magnetosphere which also produce upstream bowshocks (Leboeuf et al., 1978; Lyon et al., 1980; C. C. Wu et al., 1981). Similar progress is being made with MHD models of interplanetary flows and shock waves (Wu et al., 1979). While quite promising, simulation flow fields for a range of upstream and obstacle parameters have not yet been published and comparisons between model predictions and spacecraft observations are just beginning.

The most widely available and thoroughly tested bowshock codes are the gasdynamic models that have been used since the 1960s (Spreiter et al., 1966; Dryer and Heckman, 1967). These codes neglect the terms in the MHD equations containing the magnetic field on the grounds of their smallness in the sense  $M_a^2 = V^2/(B^2/8\pi) \sim 100$  under typical solar wind conditions. At geocentric distances less than that of the Moon, these models have been quite successful in predicting bowshock stand-off distance (Gosling et al., 1967; Fairfield, 1971; Slavin et al., 1983a) shock symmetry about the solar wind velocity direction (Slavin and Holzer, 1981; Tatrallyay et al., 1983), velocity, density, and enthalpy jumps (Burlaga and Ogilvie, 1968; Montgomery et al., 1970), magnetosheath velocity and density (Howe, 1971), and field line draping (Fairfield, 1976). Furthermore, Spreiter and Stahara (1980) have recently produced a new gasdynamic code which uses fast modern computational techniques as opposed to the old "inverse method" and allows for the specification of obstacle shape as well as upstream sonic Mach number.

Given the many successes of gasdynamic theory in describing the bowshock and flow about the magnetosphere it is important that its limitations be reviewed. The single fluid nature of these models makes it impossible for them to predict the ion and electron temperature jumps across the shock (Mihalov et al., 1980). This means that the gasdynamic models are of very limited use in predicting many important magnetosheath parameters such as  $T_e/T_i$ . In addition, the gasdynamic approximation does not hold everywhere or all of the time. As predicted by the Spreiter and Rizzi (1974) MHD model the position of the bowshock, and hence the jump conditions and flow field, becomes increasingly dependent upon  $M_a$  as Alfvénic Mach number decreases below its mean of 8-10 (Formisano, 1979a; Slavin et al., 1980; Tatrallyay et al., 1983). The plasma depletion layer at the dayside magnetopause, which may be very important to our eventual understanding of the solar wind-magnetosphere interaction processes, is an intrinsically MHD phenomena (Zwan and Wolf, 1976) and not present in the gasdynamic flow fields. The gasdynamic approximation also fails at large distances downstream where  $M_a$  in the shock frame becomes small. As would be expected for "Mach cone" arguments (Michel, 1965), the distant bowshock position has been found to be bounded by the MHD fast wave Mach cone which can be significantly larger than the gasdynamic cone angle based upon  $M_s$  (Slavin et al., 1983c), the sonic Mach number.

To summarize, fluid modeling of shock waves in space is still far from being a mature discipline. Theoretical efforts are needed to include multiple particle species, magnetic fields, time dependence, turbulence, and realistic shock and driver boundary conditions in the existing magnetohydrodynamic and gasdynamic codes. In order to accomplish these goals it is necessary that observational modeling be continued to quantitatively define the global changes in the shock and the flow field with varying upstream conditions and solar-planetary interaction mechanisms.

### Shock Extensions and Global Asymmetries

One of the key features of supercritical collisionless shocks (see below) is their ability to reflect upstream particles and transmit shock-processed downstream particles into the upstream flow. These particles excite a variety of hydromagnetic and plasma waves which in turn scatter the particles and provide a "seed" population for further acceleration between the shock and the upstream turbulence. Together the energetic particles and waves define the foreshock, an

upstream extension of the shock structure. The foreshock occurs adjacent to, and as part of, the quasi-parallel shock (Diodato et al., 1976; Greenstadt et al., 1980a); its global relationship to the bowshock can be seen in Figure 10-16 of this report.

A variety of wave types and suprathermal ion populations have been observed upstream of the Earth's bowshock which clearly are associated with the shock's nearby presence. Figure 10-4 shows schematic representations of various classes of upstream ion populations as contours of equal phase space density in 2-dimensional velocity space ( $v_x, v_y$ ). These populations include: the so-called "reflected" ions, which are nearly field-aligned beams backstreaming from the shock (e.g., Asbridge et al., 1968); "intermediate" ions, which have kidney bean-shape velocity distributions centered roughly on the field direction (Paschmann et al., 1979); "diffuse" ions, which form nearly spherical shells in velocity space (Gosling et al., 1978); and "gyrating" ions, which are phase-bunched beams having significant gyromotion about  $\underline{B}$  (Gurgiolo et al., 1981; Eastman et al., 1981). The probability of observing any of these populations depends upon a variety of factors, including distance from the shock, distance from the leading edge of the foreshock, the angle,  $\theta_{Bn}$ , between  $\underline{B}$  and the local shock normal, and the Mach number of the upstream solar wind flow. The association between various particle distributions and waves, both ULF and VLF, has been well documented. The June 1, 1981 special issue of the Journal of Geophysical Research contains a comprehensive collection of papers concerned with ISEE observations of waves and particles upstream of Earth's bowshock.

Foreshocks and associated energetic particles are observed at virtually all heliospheric shocks where in situ measurements have been possible: planetary bowshocks, interplanetary traveling shocks, and the forward and reverse shocks bounding corotating interaction regions (CIRs) in the solar wind. Furthermore, other energetic particle populations in the heliospheres are believed to originate at shocks for which in situ measurements are not yet possible: solar flare ions at coronal shocks, the anomalous cosmic ray component at the solar wind termination shock, and galactic cosmic rays at interstellar shocks. It will become apparent in later discussion that important properties of the Earth's foreshock are derived from processes involving ions reflected in the quasi-perpendicular part of the bowshock, but released to stream outward in front of the quasi-parallel shock, where their velocities are modified and they merge with the resident general foreshock ion population. Thus the global properties of nonuniform bowshocks can confound the study of the purely local properties of shocks in space.

### Local Macroscopic Shock Profiles

The preceding section treats a collisionless shock as a large-scale surface dividing two regions of plasma surrounding a stationary obstacle in a supersonic flow. But the surface is simply an envelope of shock profiles and, in fact, it is the local profiles that are actually observed, that contribute the statistics from which the surface is inferred, and that enclose the phenomenology around which shock plasma physics investigations revolve. These profiles vary with local circumstances, as the division into quasi-perpendicular and quasi-parallel sections, and description of the foreshock, have already hinted.

Figure 10-5 illustrates the extremes of shock profile as they might appear in a plot of any of the parameters of the observed plasma, such as density, temperature, magnetic field, etc.; one extreme is simple, monotonic, and stationary, the other is complex, multi-peaked, and variable. The depicted quantity, say density, increases through the shock; e.g., flow velocity, might decrease through the shock (Greenstadt et al., 1980b).

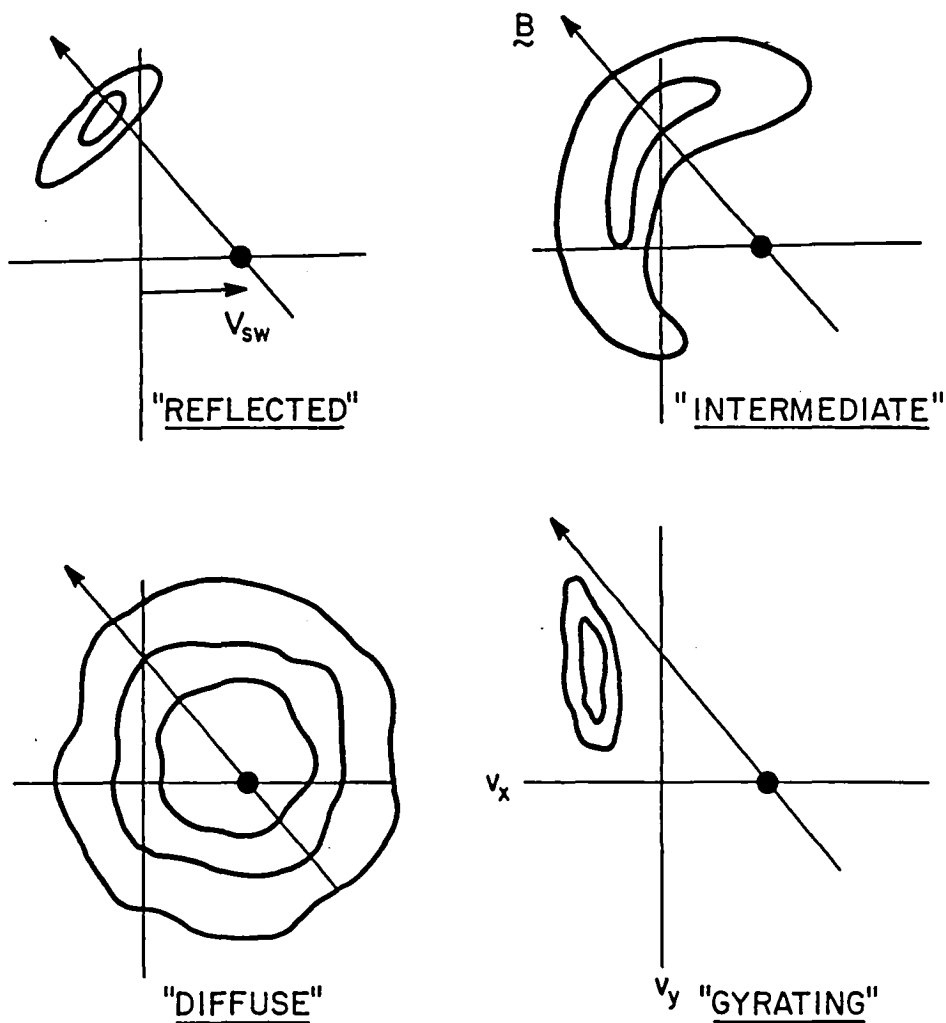
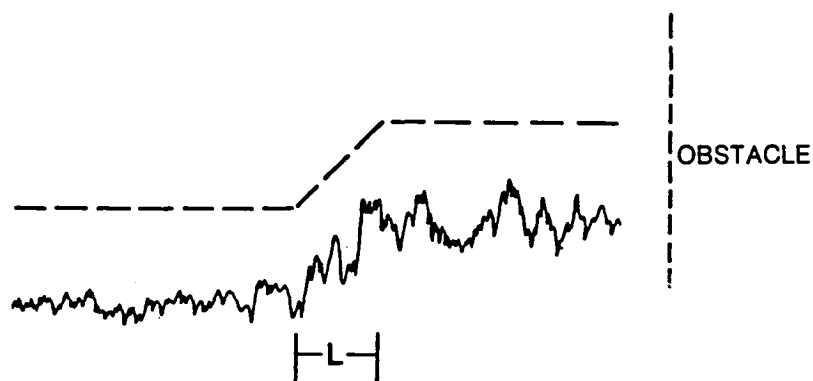


Figure 10-4. Sketches of two-dimensional contours of ion phase space distributions for four representative classes of backstreaming ions in the foreshock.



TWO EXTREMES OF SHOCK APPEARANCE

Figure 10-5. Two extremes of shock appearance, with ramp length 'L'.

The extremes may be imagined in two ways: the obvious one is that any particular plasma parameter may be thought of as exhibiting either of the depicted profiles, or some intermediate one, depending on prevailing plasma conditions. A second view, however, is that both profiles, or, more likely, less extreme versions of them, may coexist among different parameters over a range of plasma conditions. For example, the plasma speed may decline monotonically, even though the electric field may oscillate at large amplitude and rise on average. It must be borne in mind, of course, that the appearance of any parameter is procedure-dependent, i.e., with averaging over a sufficiently large interval, any variable can be made to look smooth, while with no averaging whatsoever, any variable will show some irregularity subject to the qualification that when  $\beta \ll 1$ , virtually no irregularity should be expected.

$\beta$  is, in fact, one of the major controlling variables: If  $\beta$ , in effect the ratio of thermal to field pressure, of the unshocked plasma is low, there will be little random fluctuation of plasma parameters. If  $\beta$  is very high, that is the plasma is very hot, thermal fluctuations will dominate the ambient magnetic flux; even the unshocked plasma will exhibit little or no order in any parameter, and the shock profile will be highly disturbed (Formisano et al., 1975).

A second parameter of major importance to the macroscopic features of the shock is the angle  $\theta_{Bn}$ , the angle between the ambient, preshocked magnetic field and the local normal to the shock surface. When  $\underline{B}$  is perpendicular to the local surface normal or parallel to the surface, i.e.,  $\theta_{Bn} \sim 90^\circ$ , neither charged particles nor MHD waves can travel away from the shock back into the upstream plasma, and the shock tends to have a discontinuous, or abrupt, profile. The shock is called perpendicular.

When  $\underline{B}$  is parallel to the local surface normal, or perpendicular to the surface, i.e.,  $\theta_{Bn} \sim 0^\circ$ , dispersion of both charged particles and waves becomes important: ions and waves with sufficient velocity away from the shock escape into the upstream region, modify the oncoming plasma, and spread the shock profile into a more gradual experience for at least some of the components of the approaching plasma. The shock is called parallel. It is under these circumstances that the traces of Figure 10-5 may apply to different components simultaneously, e.g., the average magnetic field increase may be spread over several Earth radii but the ion density or temperature may jump abruptly at a single transition within the extended magnetic structure.  $L$  is the distance, in either the shock or plasma frame, over which the change in average value of a component takes place. This distance is indicative of the physical process responsible for the effect on the component of interest.

In the laboratory or in theory, shocks that were neither perpendicular nor parallel were called oblique. In natural phenomenology, shocks have been designated quasi-perpendicular or quasi-parallel, with the division at about  $\theta_{Bn} \sim 45^\circ$ .

It must always be remembered that for measurements of most natural shocks in the Earth's neighborhood, the shocks are in motion, so that it is  $T_L$ , the time during which the transition occurs that is measured rather than  $L$  itself. That is, the traces drawn in Figure 10-5 represent not spatial, but a combination of spatial and temporal, variation, with the independent variable as time, not distance. Thus, the only way to obtain spatial variables of physical significance is to have an independent means of converting time to distance. This has been done in some instances, with moderate uncertainty, through various inferences, but the reliable method is to have two or more observations at locations reasonably close together, so that identifiable features can be timed as they pass from one point to another.

In view of the above remarks, the "extremes" sketched in Figure 10-5 have special significance for the study of shocks in our space environment. Multipoint observations have been the exception, rather than the rule, at least until the launch of the ISEE-1, -2 spacecraft pair. Only recently, then, have we been in a position to describe shock details quantitatively, and still only in selected cases. For example, when  $\theta_{Bn} \sim 90^\circ$ , the shock is well confined and easily identified, and examination of the records of the two spacecraft supports conversion to spatial variables. If, in addition, when  $\beta \lesssim 1$ , the shock profile is relatively simple, shock profiles can be characterized by several measurable variables. When  $\theta_{Bn} \sim 0$ , however, two spacecraft are insufficient to characterize shock profiles that are highly irregular in time and expanded over a large region of space; such shocks have been studied in a limited way, mostly by indirect or statistical techniques.

Three shock profiles are drawn in Figure 10-6, characterized by their magnetic field signatures for the purpose of defining some of the quantities accessible to paired-spacecraft observations. Each curve represents a shock typical of some combination of upstream parameters. The first, at the top, depicts a "laminar" shock, by which usage has come to mean one standing in a flow of cold plasma at low Mach number  $M$ . We shall not at this point trouble over the distinction between Alfvén and magnetosonic Mach numbers, but note that Mach number is another parameter of the upstream plasma of fundamental importance of shock physics, and to the visible structure of any given shock. In principle, the term "laminar" should apply to transitions in which all parameters change from their upstream to their downstream states without turbulence. Custom, however, has fixed laminar shocks as those for which dissipation is theoretically almost entirely resistive and principally electrons are heated, circumstances occurring when  $\beta$  and  $M$  are both low. Certainly, they are the simplest MHD shocks. Observations have established the same simplicity for low  $M$  and  $\beta$ , although heating is not confined only to electrons.

Laminar shocks exhibit the shape of the example when  $\theta_{Bn} > 60^\circ$ , but not so close to  $90^\circ$  that no waves can propagate upstream. The laminar shock of Figure 10-6, then, is described by two basic lengths and a jump. The jump,  $\Delta Q$ , is simply the net change in the average of the quantity measured, generally  $(Q + \Delta Q)/2 < 2.50$  for laminar shocks. The shock outline is determined by dispersion of a whistler mode wave whose phase velocity matches that of the solar wind so that it stands in the oncoming plasma;  $L_w$  is then the wavelength of the standing wave and the ramp thickness, the ramp being simply the last wave cycle accompanying the irreversible plasma heating of the shock. The second length,  $L_r$ , is the damping length of the standing wave in the upstream plasma; it is an inverse measure of the anomalous resistivity of the medium (Robson, 1969; Mellott and Greenstadt, 1984). For laminar shocks, the potential associated with the ramp is insufficient to reverse the directed motion of a significant number of the streaming ions; this is not the case at high flow speeds.

Still referring to quasi-perpendicular shocks, when the Mach number rises above a "critical" value, about  $M=3$ , new features appear in the shock, some of which are depicted in the center of Figure 10-6. High flow speed eliminates the possibility of standing waves ahead of the shock, when waves of no frequency can propagate across  $B$ , and the potential at the ramp becomes large enough often enough to reflect some of the streaming ions, which then execute part of a Larmor orbit in the upstream field before re-entering the shock. These recirculated particles are responsible for a "foot" of length  $L_f$  and an overshoot of length  $L_o$  before and behind the main ramp, respectively, and form a "secondary" population of ions whose energy contributes to the initial dissipation of the bulk flow energy by dispersing some of the bulk velocity and which is ultimately thermalized by the shock along with the bulk population. The two populations eventually blend into a single, stable downstream distribution many Larmor radii behind the ramp, in a sense defining one measure of supercritical shocks, an overall ion thermalization length  $L_i$ . The overshoot  $O$  tends to increase with increased  $M$ , and the ramp itself has a length  $L_r$  on the order of the ion inertial length  $c/\omega_{pi}$  rather than a dispersive length (Leroy et al., 1982). High frequency waves are always visible in addition to the other features.

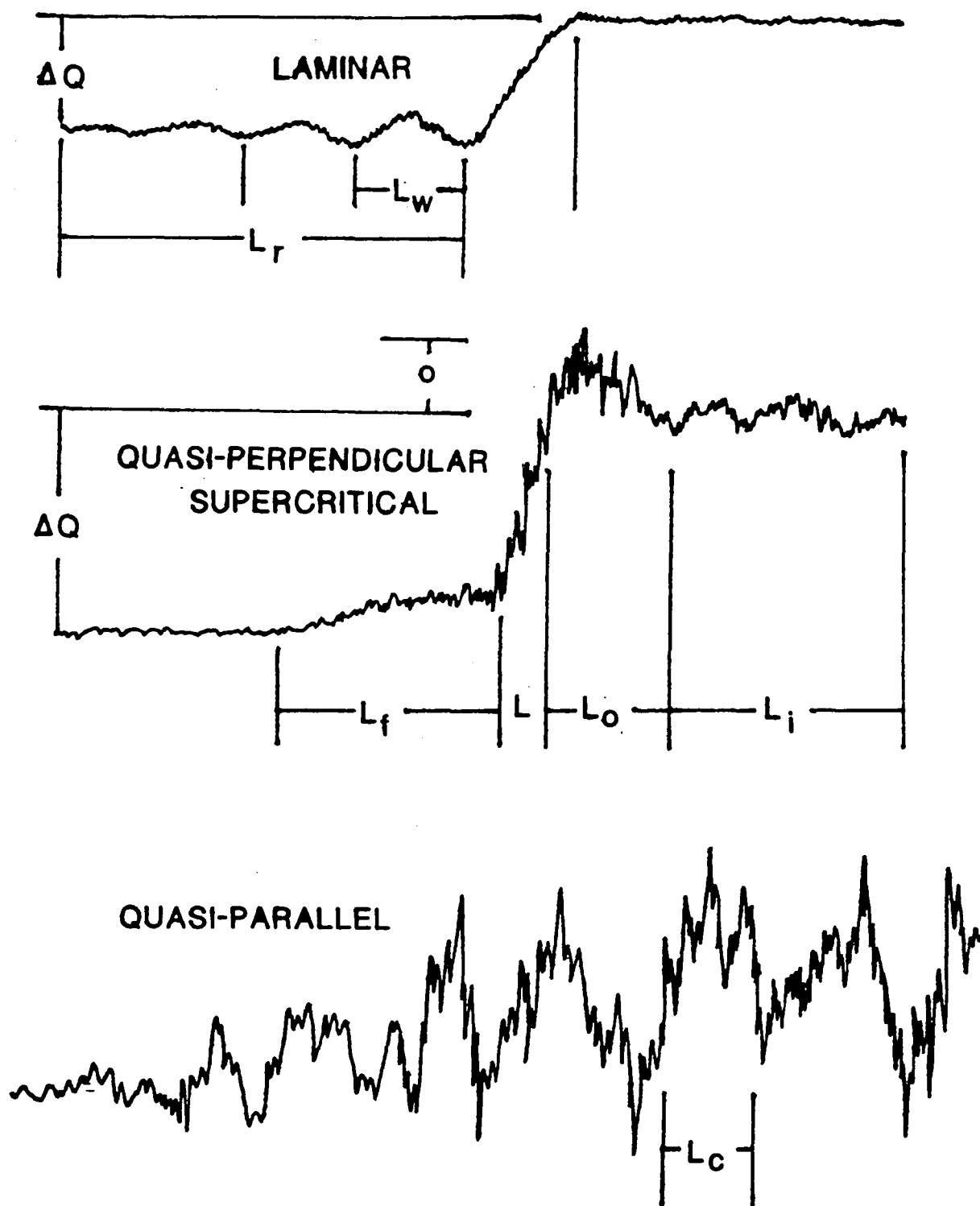


Figure 10-6. Sketches of three representative shock profiles, showing various thickness parameters defined in the text; upstream is to the left.

Ions of the secondary population generally have nonzero velocity components along the ambient magnetic field, i.e., the guiding centers of these ions also move along  $\underline{B}$ . When  $\theta_{Bn} < 45^\circ$ , some of the circulating ions can actually move upstream away from the shock and return to the unshocked plasma, traveling upstream into the unshocked plasma far ahead of the shock and spread out into the upstream plasma, interact with it, modify it, and create waves that are carried into the shock to modify its structure as well. The result is a thick, complicated foot, or "foreshock", and a thick, complicated shock whose detailed physical properties become difficult to investigate in relation to the upstream parameters because of the difficulty of capturing the appropriate average values. To some instruments, the transition between foreshock and shock can become quite indistinct. The bottom trace in Figure 10-6 is meant to illustrate this shock form, which has been studied little and only statistically, yielding a correlation length  $L_c \sim 1000$  km within which the largest magnetic perturbations retain a quantitatively identifiable shape (Greenstadt, et al., 1982).

The most common means of investigating the quasi-parallel form has been by close examination of the wave and ion properties of the foreshock, where the primary and secondary ion populations are distinguishable, and the waves are relatively small.

The foregoing paragraphs provide an oversimplified discussion of local macroscopic shock structure intended as observational background for the more specific details described in later sections. Basically, shock profiles divide themselves into two broad categories, according to whether reflected ion components can or cannot be confined to a narrow transition from upstream to downstream regions: Shocks are quasi-perpendicular when roughly,  $\theta_{Bn} > 45^\circ$ ; shocks are quasi-parallel when, roughly,  $\theta_{Bn} < 45^\circ$ . Within the quasi-perpendicular category, the relative ease with which the shock can be discerned, and its upstream parameters determined, has enabled us to define subcategories of structure: Laminar for cold, low Mach flow; quasi-laminar for cold, supersonic (high Mach) flow; quasi-turbulent for hot, low Mach flow; turbulent for hot, supersonic (high Mach) flow (Formisano and Hedgecock, 1973a). Figure 10-7 is a three-dimensional conceptualization showing those blocks of  $\beta$ -Mach- $\theta_{Bn}$  space that have been the subjects of detailed study of observations, theory, and simulation. Clearly, most of the volume of this parameter space is still open to investigation, although it should not be necessary to examine every corner with equal care if key parameter combinations are judiciously chosen. We seek mainly to establish the basic classes and subclasses in which shock types fit and, of course, to understand why.

## KEY QUESTIONS

The central question in all investigations, experimental or analytical, natural or contrived, has been, and remains: What are the collisionless processes whereby supersonic ion streaming energy is transformed into thermal and suprathermal energy? Interest in this question is sharpened by observations that high Mach number collisionless shocks heat ions appreciably more than they heat the much lighter electrons of the original upstream plasma.

Two decades of experience with shocks in space have added to the above some questions that might also be considered "key" in the sense that they motivate the studies encompassing many more detailed issues. One such question is: How is a shock modified by being nonuniform? That is, if an extended shock "surface" consists of more than a single, uniform structure, how do the local structures affect each other? A second question: What do shocks feed back into their overall environment? In other words, what role do shocks, their foreshocks, and their products contribute to the general cosmological scene?

Considerable progress in answering these questions is evident in some areas, as the following sections reveal.

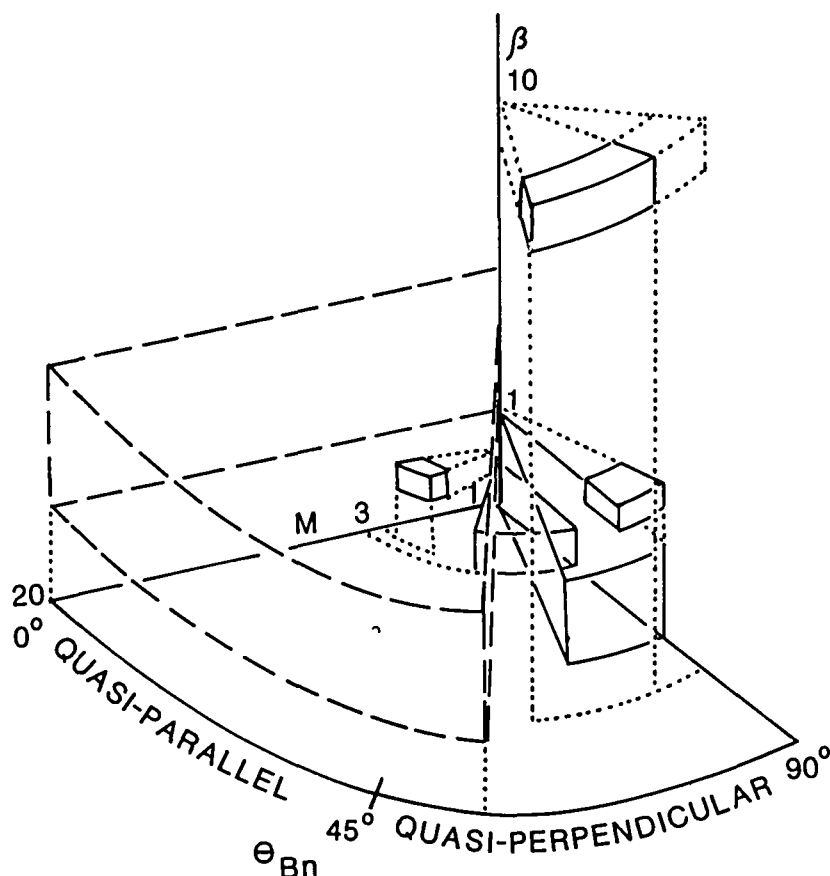


Figure 10-7. Regions of  $M$ - $\theta_{Bn}$ - $\beta$  space that have been intensively studied to date; dashed wedge at left is to indicate that the quasi-parallel shock has been investigated, but mainly indirectly by examination of the foreshock.

### III. QUASI-PERPENDICULAR, SUPERCRITICAL PROCESSES

In this section we describe the magnetic, particle, and wave structure of the quasi-perpendicular, supercritical shock as presently constructed from laboratory, theoretical, and spacecraft sources. We draw heavily on the most recent model developed from the simulation results (Leroy et al., 1981, 1982) obtained using a hybrid code (see Section VIII), in combination with bow-shock observations, whose recent high resolution, multipoint measurements have guided the simulations and validated much of their application to natural space plasmas. Laboratory results are necessarily historical in tone, since few, if any, experiments have been conducted in recent years, but the early results are still of importance.



## ION DYNAMICS AND FIELD STRUCTURE

### Laboratory-Shock Structure

During the late 1960s through the early 1970s, a number of laboratory experiments were performed to study collisionless shock waves. The shock waves were produced either by driving a magnetic piston into a stationary plasma, or by directing a super-Alfvénic plasma flow against a stationary magnetic obstacle. These experiments produced a great deal of detailed information on collisionless shocks. However, their relevance to geophysical shocks is somewhat limited due to practical constraints on both classes of experiments. Although the piston experiments could produce shocks for a wide range of parameters, their duration was limited by the motion of the shock in the apparatus to about one ion gyroperiod or less. The plasma stream experiments were much longer in duration, but were performed only over a very restricted parameter range. A further difficulty in both types of experiment was the establishment of collisionless conditions upstream of the shock.

Although shocks were observed incidentally in a number of  $\theta$ -pinch and z-pinch fusion experiments, the most detailed laboratory information came from experiments specifically designed for shock generation. Two experiments in particular yielded a wealth of information: these are the Tarantula experiment at Culham Laboratory (Paul et al., 1971), and the oblique shock experiment at the University of Texas (UT): (Robson and Sheffield, 1969). Well separated, quasi-stationary shocks were produced with  $M_A$  2-5. In Tarantula, the shocks were perpendicular to the applied field; in the UT experiment, a curved shock front was created in which the angle of the shock to the field varied from  $90^\circ$  at mid-plane to about  $40^\circ$  at the edges. Since the radius of curvature of the shock front was large compared with the shock thickness, local regions of the shock front were treated as plane. The  $90^\circ$  shocks in the UT experiment and Tarantula were identical.

Above  $M_A=3$  the perpendicular shock showed a characteristic magnetic precursor (the "foot"), and at the same time the electron heating was found to account for less than the available dissipation in the shock. It was speculated for some time that the foot was caused by reflected ions, and these were eventually detected using special probes (Phillips and Robson, 1972). Downstream effects of the reflected ions were not determined; no ion temperature measurements were taken, although the reflected ions could reach the downstream region during the limited duration of the experiment. A stationary periodic magnetic structure of small amplitude was sometimes seen behind the main shock transition. As  $M_A$  was increased the number of reflected ions increased, but the potential jump across the shock, normalized to the flow energy, decreased. Thus the ions were not reflected by the main potential jump. It was suggested that the reflection was caused by an electrostatic ion-acoustic wavetrain which followed the main jump, but such a wavetrain was not actually detected and in any case would have been too small to be resolved by the electric probes.

Experiments at UT showed that the structure of the oblique shock was quite different from that of the perpendicular shock. Theory makes a fine distinction between truly perpendicular shocks and "oblique" shocks whose normal angle  $\theta_{BN}$  is, roughly, less than  $89^\circ$ ; specifically, less than  $(m_e/m_i)^{1/2}$  radians (Tidman and Krall, 1971). In oblique shocks, circularly-polarized wavetrains of scale length  $c/\omega_{pi}$  extended ahead of the main density jump and were identified as large-amplitude whistler waves stationary in the shock frame. Shortly after the formation of the whistler, a shorter wavelength disturbance was seen to grow on the whistler structure (Decker and Robson, 1972). This had a scale of  $0.2c/\omega_{pi}$  and was tentatively identified as a decay instability. The magnetic structure changed in less than an ion transit time, so ions could have been scattered

and thermalized by this instability. To account for the damping of the whistler it was necessary to introduce an anomalous resistivity, but no measurements of the turbulent density fluctuations were made in these shocks. Measurements of the electron temperature behind the shock were consistent with purely electron heating up to the critical  $M_A$  (which varies with angle), were for higher  $M_A$  an increasing fraction of the shock energy was not accounted for, and was presumed to be going into the ions, although no reflected ion measurements were made in the oblique case. In another experiment, a curved shock front was formed which showed the characteristic whistler structure when it was inclined locally at  $20^\circ$  to the field ( $\theta_{Bn} = 70^\circ$ ).

The most detailed observations of shocks using the plasma-stream, rather than transient-field, approach were made in a plasma wind-tunnel at Cornell University (Morse and Destler, 1972). The Mach number, based on the fast magnetosonic speed, was 3-5 and the shock was predominantly perpendicular. This experiment had the advantage of large size, and produced a standing shock whose thickness was  $c/\omega_{pi}$ ; the downstream electron temperature accounted for 40% of the available shock heating, but no measurements were made of the ion temperature. Strong magnetic fluctuations were measured at the ion cyclotron frequency and the position of the shock was observed to oscillate in the direction of the flow, also at about the upstream ion gyrofrequency. In contrast to the piston experiments, no potential jump was found across the shock.

### Natural Bowshock of the Earth

Spacecraft observations of the Earth's bowshock have also indicated the significance of reflected ions, with nonthermal ion populations observed both up and downstream of the shock. A second peak in the ion velocity distribution in the foot and just downstream of the shock was reported by several observers (Montgomery et al., 1970; Formisano and Hedgecock, 1973b; Bame et al., 1979; Greenstadt et al., 1980b). Solar wind ions are essentially unmagnetized (i.e., do not conserve their magnetic moments) in the bowshock transition because the main jumps in magnetic field and electrostatic potential at the shock occur over a distance considerably shorter than an ion gyroradius (e.g., Russell and Greenstadt, 1979). Thus, single reflections of solar wind ions at the Earth's bowshock should be nearly specular (Schwartz et al., 1983); that is, some of the ions can simply behave approximately as if they reversed their velocity component normal to the shock while leaving their tangential component relatively unchanged. Recently, some of the ions of the secondary peaks have been characterized as "gyrating ions" and identified as ions specularly reflected at the shock (Paschmann et al., 1982).

In space, it has been virtually impossible to certify any shock as exactly perpendicular within the uncertainties of measurement, but in any case, reflected ions are detected in quasi-perpendicular shock profiles over a wide range of "oblique" shock-normal angles (Schopke et al., 1983). Moreover, secondary ion profiles appear much the same for all observed quasi-perpendicular shocks. Since the Mach number of the solar wind flow at 1 AU is usually within the range 4-10, the bowshock is supercritical most of the time, and ion reflection is a characteristic feature of the bowshock. This phenomenon will be discussed at length below.

The significance of ion reflection is that it produces scattering in velocity space and thus constitutes an important step in the process by which the shock converts ion streaming energy into thermal energy. Spacecraft observations have provided evidence that reflected ions account for the bulk of the heating at quasi-perpendicular, supercritical shocks (Montgomery et al., 1970; Schopke et al., 1983).

As far as oblique whistler structure is concerned, the bowshock has exhibited stationary whistler structure only under unambiguously oblique conditions ( $\theta_{Bn} < 89^\circ$ ) and only at low Mach number. Further discussion of this phenomenon is therefore included in the separate sections on subcritical shocks. At higher Mach numbers, propagating whistlers of VLF frequencies are a well known phenomenon in and around the shock (Heppner et al., 1967; Smith et al., 1967; Guha et al., 1972; Fairfield, 1974; Rodriguez and Gurnett, 1975; Greenstadt et al., 1981), but these waves do not play a major role in the energy budget of the shock.

### Numerical Simulations

The reflected ions detected in the laboratory, and thought to be responsible for the foot of the shock, were also prominent in computer simulations performed during the period of laboratory shock interest (Auer et al. 1962; Biskamp and Welter, 1972a; Papadopoulos et al., 1971; Forslund and Friedberg, 1971; Mason, 1972; Chodura, 1975). However, as in the laboratory, the relevance of the computer models to geophysical shocks was limited. Generally, the simulation codes used considered only one spatial dimension and did not include resistivity due to wave turbulence. Most particle simulations neglected the magnetic forces on the ions (valid for times less than an ion gyroperiod, i.e., piston experiments) and found highly time-dependent structure for which the reflected ions escape upstream (Papadopoulos et al., 1971; Forslund and Friedberg, 1971; Mason, 1972). Auer et al. (1962), Auer and Evers (1971), and Biskamp and Welter (1972a) treated magnetized ions, and thus could study the effects of the eventual passage of reflected ions into the downstream region. They also found strong time dependence; however, resistivity was not included and effects of the piston on the ion trajectories could not be neglected in this work. Chodura (1975), who included resistivity in a phenomenological manner but considered unmagnetized ions, found that resistivity limited the number of ions reflected.

Computer simulations (e.g., Forslund and Friedberg, 1971; Biskamp and Welter, 1972a; Leroy et al., 1981, 1982; Quest et al., 1983) indicate that shocks reflect an increasing fraction of the incident solar wind ions as the Mach number increases, with appreciable enhancement above a critical value  $M_{\tilde{C}} \sim 2-3$ , and the latest simulations have confirmed the consistent appearance and importance of the reflected ions, which constitute the key dissipation process in the shock model of Leroy et al. (1981, 1982).

The numerical model of Leroy et al. is illustrated in Figure 10-8. In the background of the figure the typical magnetic field structure is shown. Obtained from a hybrid simulation (Leroy et al., 1981), the field exhibits the features characteristic of the observed structure (Heppner et al., 1967; Russell and Greenstadt, 1979; Bame et al., 1979; Russell et al., 1981). These features (see Figure 10-6, center) are the "foot" - the region of gently increasing field strength, the "ramp" - the region marked by steep increase in field strength, and the "overshoot" - a region following the ramp where the field exceeds its equilibrium downstream value (which satisfies the Rankin-Hugoniot relations). The overshoot is followed by an "undershoot"; in fact the field can make several excursions around its downstream value before relaxing to this value. Ion behavior is shown in the foreground of the figure where the typical ion trajectories in the velocity plane ( $v_x - v_y$ ) perpendicular to  $\underline{B}$  are shown as functions of distance along the shock normal  $n$ . For simplicity, the figure has been drawn for a perpendicular shock; however, the general model is valid for  $\theta_{Bn} > 50^\circ$ ; i.e., both the "oblique" shock of laboratory prominence and the pure perpendicular shock exhibit essentially the same "quasi-perpendicular" field and ion reflection characteristics.

Basically, the ion behavior is the following: The upstream ions approach the shock with super-Alfvénic normal velocity  $v_n$ . In the foot region they are magnetically deflected toward  $-v_y$  while decreasing  $v_n$ . Near the peak at  $B$ , at the maximum in the electric potential, some ions are reflected and flow back into the foot region; the majority of the ions pass directly into the downstream region. The reflected ions are magnetically deflected in the upstream magnetic field,

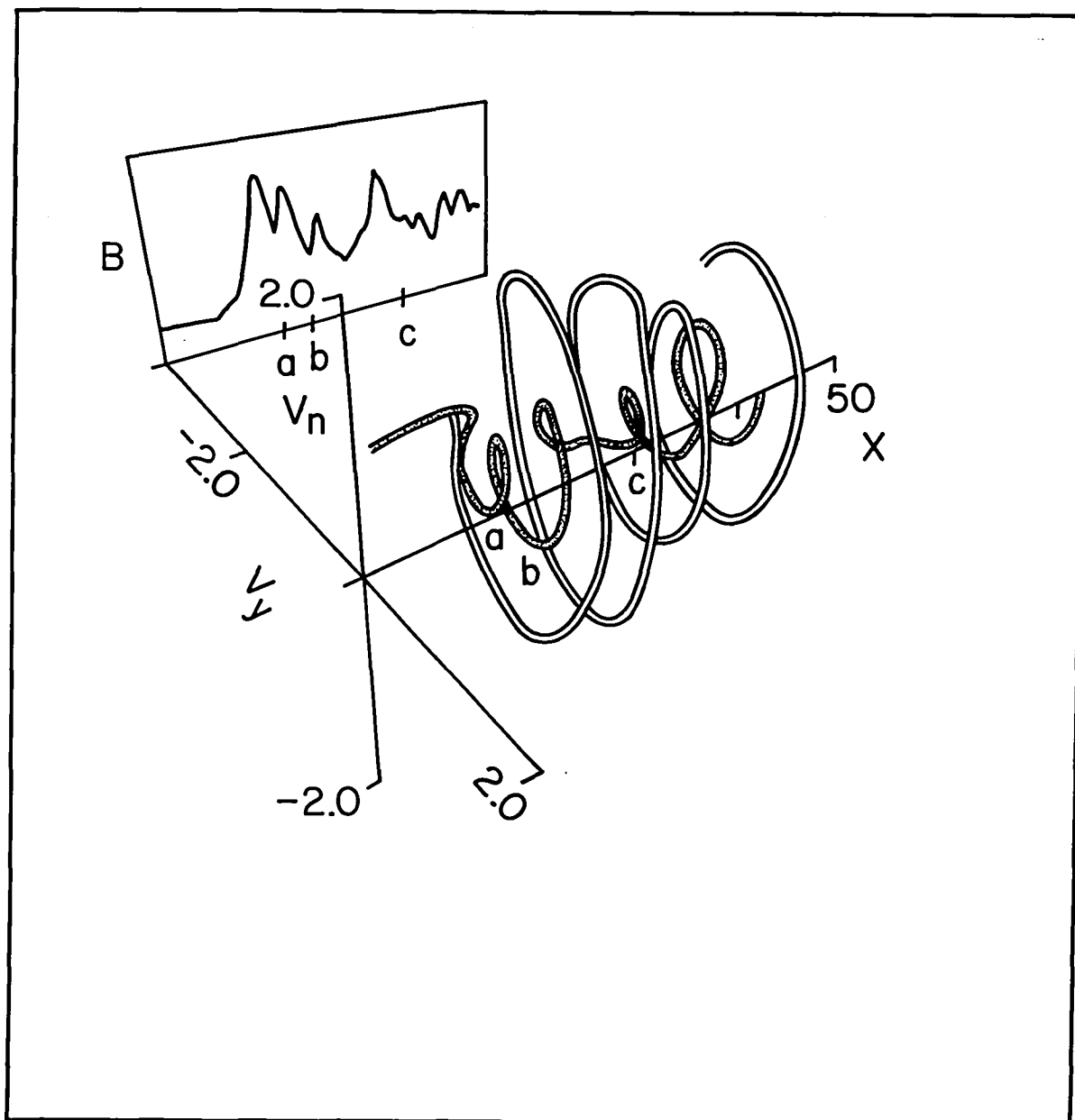


Figure 10-8. Schematic diagram of a perpendicular shock wave. Shown in the foreground are the  $(v_n - v_y)$  trajectories of the directly transmitted (shaded) and initially reflected (unshaded) plasma components through the shock. The shock's profile is shown in the background of the figure. The points  $a$ ,  $b$ ,  $c$  indicate the position of the peaks in  $B$  along the plasma trajectory.

reverse direction and subsequently enter the downstream region. Since the velocity of these ions is much larger than that of the directly transmitted ions, both ion populations have helical trajectories downstream around the center of mass velocity.

As shown in Figure 10-8, the magnetic field structure is closely related to the ion trajectories. In fact, with the exception of the ramp, the magnetic field features are a result of the reflected ion gyromotion and thus have scale lengths related to the gyroradius  $P_r$  of these ions as discussed in detail in Leroy et al., (1981, 1982). This can be understood most easily by noting that for nearly-perpendicular shocks, assuming that the resistivity is reasonably small outside the ramp, the magnetic field magnitude is approximately proportional to the density (e.g., Olbert and Rossi, 1970). The extension of the reflected ion trajectories into the upstream region, which increases the density over approximately  $0.7 P_r$ , causes the foot. The overshoot results from two effects. First, although the bulk of the incoming ions penetrate directly into the downstream region, they are decelerated to normal velocities significantly smaller than the asymptotic value  $V_2$ . This causes a large density increase directly behind the ramp. The second effect is that  $P_r$  is large enough that after entering the downstream region, the reflected ions return close to the ramp, increasing further the density and hence the magnetic field in the overshoot region. Subsequent density and magnetic field bunching occurs at points, separated by  $P_r$ , where the normal velocities of the directly transmitted and reflected ions are small. These features extend downstream until some process, such as phase mixing or wave turbulence scattering, thermalizes the ion distribution.

The ramp itself is not determined by the ion gyromotion, but rather by the electron motion. An electric field  $E_n$  along the shock normal exists in the ramp. The role of this electric field in decelerating and reflecting the incoming ions and in accelerating the electrons is discussed in some detail below. Electron phenomenology is not treated by the simulation code discussed above, because electrons are treated as a fluid there. Our direct knowledge of electron distributions and their evolution through the shock comes from spacecraft measurements. These are described in a later subsection.

### Comparison of Simulation with the Observations of 7 November 77

A result from the model of the supercritical quasi-perpendicular shock is illustrated in Figure 10-9. In Figure 10-9a is the ISEE-2 magnetic field profile for the bowshock crossing of November 7, 1977 (Sckopke et al., 1983). This was an almost perpendicular shock ( $\theta_{Bn} \sim 82^\circ$ ) with  $M_A = 8$ . The dashed line in the figure is the simulation field obtained using the hybrid code with the observationally determined parameters. The comparison was possible because the speed of the shock could be determined by combining the ISEE-1 and ISEE-2 data, thus allowing the time scale to be converted to a spatial scale. The agreement of the observed and simulated structure indicates that the shock field is dominated by ion gyromotion effects. The spatial extent of the agreement further indicates that ion thermalization, which would eliminate the gyromotion, occurs with a scale length much larger than  $P_r$ .

The accord of the observational and simulation results extends to the level of the ion velocity space distributions. A sequence of ion measurements taken by the ISEE-1 FPE experiment during the November 7, 1977 bowshock crossing is presented at the top of Figure 10-9b (Sckopke et al., 1983). Shown are ion isodensity contours in the velocity space perpendicular to the ISEE-1 spin axis. The upstream field  $\underline{B}$  is inclined roughly  $45^\circ$  to this plane. The timing of each measurement is indicated by a dot on the magnetic field profile in the plot inserts. In the rightmost plot, taken at the upstream edge of the foot, the beam of reflected ions is apparent in addition to the cold solar wind beam. Progressing into the shock (to the left) the reflected ions become more pronounced and disperse. At the ramp and beyond, complex combinations of the reflected ions and decelerated solar wind are seen. Equivalent representations of the ion distribution functions from the simulation results are shown at the bottom of Figure 10-9b. A numerical model of the

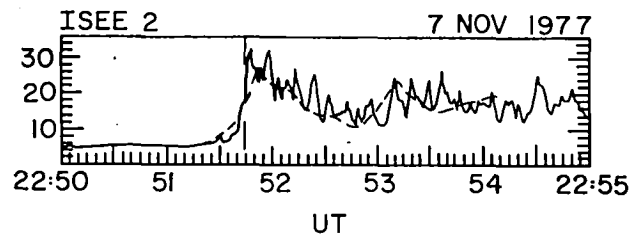


Figure 10-9a. ISEE-2-profile for the quasi-perpendicular bowshock crossing of November 7, 1977. The dashed line shows the simulation results for this event plotted as an equivalent time sequence;

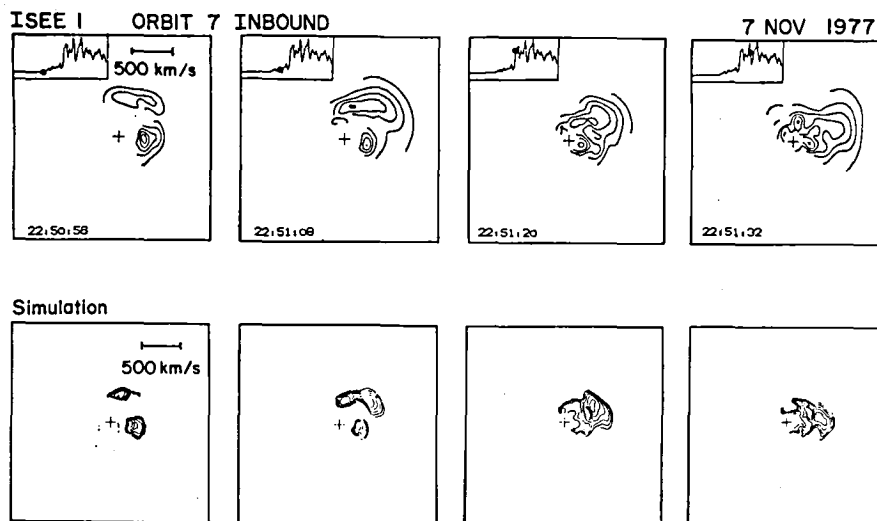


Figure 10-9b. Isodensity contours of the positive ion distribution function for the November 7, 1977 crossing. The upper contours show ISEE-1 results from the FPE experiment (Schopke et al., 1983). The position of each measurement along the profile of B is indicated in the insert. The lower contours were obtained at equivalent positions in the hybrid simulation.

FPE instrument which incorporates the energy, time, and geometric resolution of the FPE instrument, was used as a diagnostic in the hybrid code. In the observed distributions, the reflected ion beams are somewhat less sharply defined. This probably indicates a small amount of scattering due to wave turbulence present, a possibility considered in more detail below. (Some of the difference can be attributed to somewhat different contour levels and instrumental noise.) However, the overall agreement is encouraging.

### Ion Reflection Machinery

A central question, deferred above, is that of the actual ion reflection mechanism. This question has been illuminated by a detailed examination of the ion motion in the hybrid simulations (Leroy et al., 1981, 1982). The results show ion reflection to be a self-sustaining process in which reflected ions in the foot region essentially produce further reflected ions.

The presence of reflected ions in the foot region facilitates further ion reflection in two ways: first, they significantly increase the maximum shock electric potential  $\phi$  and second, they cause the incoming ions to be magnetically deflected in the foot. The electric potential  $\phi$ , the negative integral of the shock electrostatic electric field  $E_n$ , can be expressed in the limit of small electron mass as

$$e\phi = \int_{-\infty}^{\chi} \frac{1}{n} \frac{d}{dx'} \left( \frac{B^2}{8\pi} + P_e \right) dx' + \frac{1}{c} \int_{-\infty}^{\chi} v_y B_z dx' \quad (1)$$

Here  $P_e$  is the electron pressure. The first integral term on the RHS,  $\phi$ , represents the potential change due to the pressure increase in the shock and is important for subcritical as well as supercritical shocks. The second integral,  $\phi$ ; is the contribution due to a bulk drift of the ions perpendicular to both  $\underline{n}$  and  $\underline{B}$  which accompanies ion reflection. This term is quite significant for supercritical shocks, as pointed out by Auer et al. (1971). The potentials are shown at the top of Figure 10-10 for the simulation of a perpendicular shock with  $M_A=8$  (Leroy et al., 1981). The contribution of  $\phi$ , shown shaded in the figure, results in a characteristic overshoot structure thinner than the magnetic field overshoot. The overshoot is formed as follows. In the foot region, the bulk ion cross field drift  $V_y$  is dominated by the reflected ions making  $V_y$  positive, as shown at the bottom of Figure 10-10. At, and somewhat behind, the ramp  $V_y$  is dominated by the transmitted ions and is negative, decreasing  $\Delta\phi$ .

The potential overshoot resulting from the reflected ions has two implications. The value of  $\phi_{\max}$  is increased, more than doubled for the case shown here. This increase facilitates ion reflections, but is still too small to reflect ions alone;  $e\phi_{\max}/1/2mv^2 \sim .7$  here and actually decreases with increasing  $M_A$  (Leroy et al., 1982). In addition an electrostatic barrier or "plug" inhibits upstream leakage of the downstream ions.

The additional deceleration along the shock normal of the incoming ions required in the foot is accomplished as follows. The foot acts as a magnetic brake where the incoming upstream ions are magnetically deflected, decreasing their normal velocity enough for some ions to be reflected by the potential. To understand this process, consider the equations of motion of an incoming ion, given here for simplicity for a perpendicular shock:

$$\frac{d}{dt} V_n = - \frac{e}{M} \frac{d}{dx} \phi - \frac{e}{Mc} V_y B_z \quad (2)$$

$$\frac{d}{dt} V_y = - \frac{e}{Mc} (v_n - V_n) B_z \quad (3)$$

Here  $V_n$  and  $V_y$  are the individual ion velocities, while  $v_n$  is the bulk ion normal velocity. We have approximated  $E_y$  by  $1/c v_n B_z$  in Equation (3) which is valid in the foot. Because of the presence of reflected ions in the foot,  $v_n$  is smaller than  $V_n$  for the incoming ions, which, according to Equation (3) are deflected to negative  $V_y$ . This deflection decreases  $V_n$  (Equation 2).

In summary, the reflection and subsequent transport downstream of a fraction (10-20%) of the upstream ions is found to provide a key process in determining the overall structure of supercritical quasi-perpendicular shocks. Important characteristics of the shock include:

- Ion reflection - ions are reflected at the ramp by a combination of electrostatic and magnetic forces which depend critically on the simultaneous presence of reflected ions in the foot region. In the foot, in front of the shock surface, the reflected ions have a large ( $\gg$  thermal speed) velocity relative to the incident solar wind ions and electrons.
- Ion heating - the large velocity difference between the gyrating and core ions may excite instabilities that produce the eventual downstream jump in ion temperature.
- Electric potential structure - the gyrating ions cause a net ion drift velocity perpendicular to both the shock normal and the magnetic field. This drift greatly increases the potential in a narrow region centered at the ramp, resulting in an "electrostatic plug" that isolates the downstream ions from the upstream region. In addition, this electric field confines electrons within the downstream region to neutralize shocked ions.

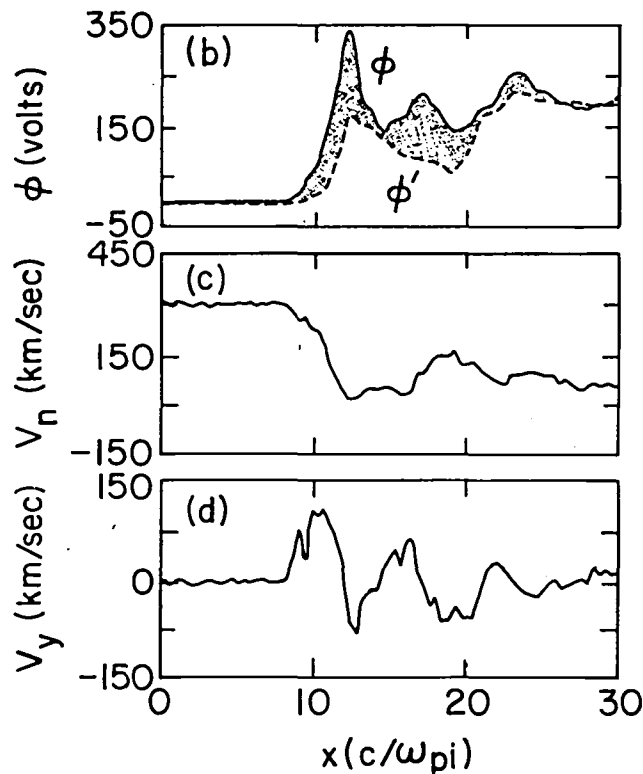


Figure 10-10. Profiles of the simulation results for  $\theta$ ,  $v_m$ , and  $v_y$ .



- Magnetic field and density overshoots - the gyrophase bunching of both the gyrating and core ions leads to density and associated magnetic field bunching which cause the overshoot structure.

## ELECTRONS

The evolution of the electron distribution across a collisionless shock is determined by the combined effects of the macroscopic (mainly d.c.) magnetic and electric fields and the microscopic (fluctuating, short wavelength) fields at the shock. The macroscopic fields are responsible for a reversible gain in the electron energy across the shock, whereas the microscopic fields lead to an irreversible increase in the random energy of the distribution.

As described in the previous section, the macroscopic field configuration of the shock consists of a rapid jump in the magnetic field strength, accompanied by an outwardly directed electrostatic field which is required to help slow the bulk plasma flow. In a supercritical, quasi-perpendicular shock, the electrons are related to the field configuration as follows.

The change in magnetic field strength at the shock is explained by the presence of a current perpendicular to the magnetic field which is carried by a net perpendicular drift of the electrons relative to the ions. This drift arises because the electron gyroradii are small compared to the ramp thickness, so that they undergo a drift in  $E_{\perp}$  induced by  $\underline{E} \times \underline{B}$ , unlike the unmagnetized ions. The drift has been directly observed in a quasi-perpendicular shock by Scudder et al. (1983). Moreover, the cross-field current density implied by the magnetic field jump has recently been determined from ion and electron measurements (Scudder et al., 1983) for a well-studied bow-shock crossing. This current density and the associated change in  $\underline{B}$  confirm that the thickness of the cross-field current-carrying layer is significantly thicker than the electron gyroradius, but less than the convected ion gyroradius.

Since the shock is thick compared to an electron gyroradius and the transit time for the plasma to cross the shock is long compared to an electron gyroperiod, electrons should conserve their magnetic moments in the transition, gaining gyration energy as they traverse the jump in magnetic field strength. The perpendicular electron temperature should thus increase directly proportional to  $B$ , and could consequently exceed the parallel electron temperature. The quantity  $T/B$  has been found to be nearly conserved across the shock (Scudder et al., 1983; Feldman et al., 1983b), and at least for the case of weaker shocks, simple conservation of the magnetic moment provides a good mapping of the upstream to the downstream perpendicular velocity distribution for accessible orbits (Feldman et al., 1983b).

The macroscopic electrostatic field of the shock is parallel to the shock normal and thus, except in exactly perpendicular shocks, has a component parallel to the magnetic field,  $\underline{B}$ . One effect of this electric field should be to accelerate incident electrons into the downstream region in the direction parallel to  $\underline{B}$ . In fact, the electron distribution function within the bowshock transition is frequently observed to exhibit a peak in the parallel direction centered roughly on the downstream edge of the developing flat top (Feldman et al., 1982, 1983a); the distribution will be discussed below. With deeper penetration into the shock from the upstream side, the height of the offset peak diminishes, the offset speed increases (as does the speed at the edge of the flat top), and the effective temperature of the peak decreases. The  $\theta_{Bn}$  dependence of the occurrence of such field-aligned peaks has not yet been investigated. It has been suggested (Feldman et al., 1982, 1983a) that this peak in the parallel direction is the result of the acceleration of the incident solar wind electron distribution by the field-aligned part of the shock electrostatic field.

This suggestion needs further quantitative investigation, but if it is correct, then the speed of the beam and hence the break-point of the downstream flat top could be a measure of the parallel potential jump experienced by the electrons.

Observed changes in the electron distribution appear to depend in a general way on the strength of the shock, (Feldman et al., 1983b). For weak shocks, electron heating occurs almost exclusively in the direction perpendicular to the magnetic field. As the shock strength increases, heating in the parallel direction becomes more apparent. At the same time, the downstream distribution begins to exhibit an increasingly non-Maxwellian appearance, characterized by a nearly isotropic flat top at low energies. This flat top is a well known feature of electron distributions downstream of the Earth's bowshock, (e.g., Montgomery et al., 1970; Scudder et al., 1973; Feldman et al., 1983a), which usually exemplifies a relatively strong collisionless shock (Greenstadt et al., 1980b). When the flat top is present, it is often found that the downstream parallel electron temperature exceeds the perpendicular temperature, implying that significant electron energy gain has resulted from traversal of the electrical potential of the bowshock. Figure 10-11 shows two typical examples of the observed upstream to downstream change in the electron distribution (after Feldman et al., 1983b). The upper set of panels show the upstream and downstream distributions in the parallel and perpendicular directions for a relatively weak interplanetary shock observed on 4 October 1978. The lower two panels show the same distributions for a relatively strong interplanetary shock observed on 24 April 1979.

The most striking feature of the lower sets of cuts (c and d) is the formation of a flat top at low energies in the downstream distribution. The open circles in Figure 10-11d show the downstream perpendicular distribution one would expect assuming that the magnetic moment of each electron is conserved and that there is no potential jump. The shape of the observed distribution, especially the flat top at low energies, is incompatible with conservation of magnetic moment as the only important effect.

The break point energy of the electrons' flat top is rarely observed to exceed  $\sim 150$  eV, whereas the electrostatic potential jump is generally regarded as being a significant fraction ( $\sim 0.5$ ) of the upstream ion bulk flow energy per charge, i.e., several hundred volts. The electrons do not experience the entire potential jump across the shock. This effect is probably due to the fact that the electrons gain energy only to the extent that their motion parallel to the local  $B$  through the layer is parallel to the total (convective plus shock electrostatic) electric field within the shock. The  $E \times B$  drift motion of the electrons allows them to convect across the shock without crossing as many equipotential contours as they would if they were completely unmagnetized (Feldman et al., 1983a; Scudder and Goodrich, 1983).

The electric and magnetic field structure, with its electron acceleration, drift current, and moment conservation, creates a background of non-Maxwellian distributions of electrons which interact with one another and with the subpopulations of protons to generate wave-instabilities that contribute to both electron and ion thermalization.

## INSTABILITIES AND WAVES

Several plasma instabilities suggest themselves in light of the expected and observed changes in the ion and electron distribution functions enumerated above. The perpendicular heating associated with magnetic moment conservation is a potential source of whistler waves (Tokar et al., 1983; Wu et al., 1983b). The cross-field current could potentially drive Buneman, ion acoustic, or lower-hybrid-like modes (e.g., Lemons and Gary, 1978; Wu et al., 1983b). The acceleration

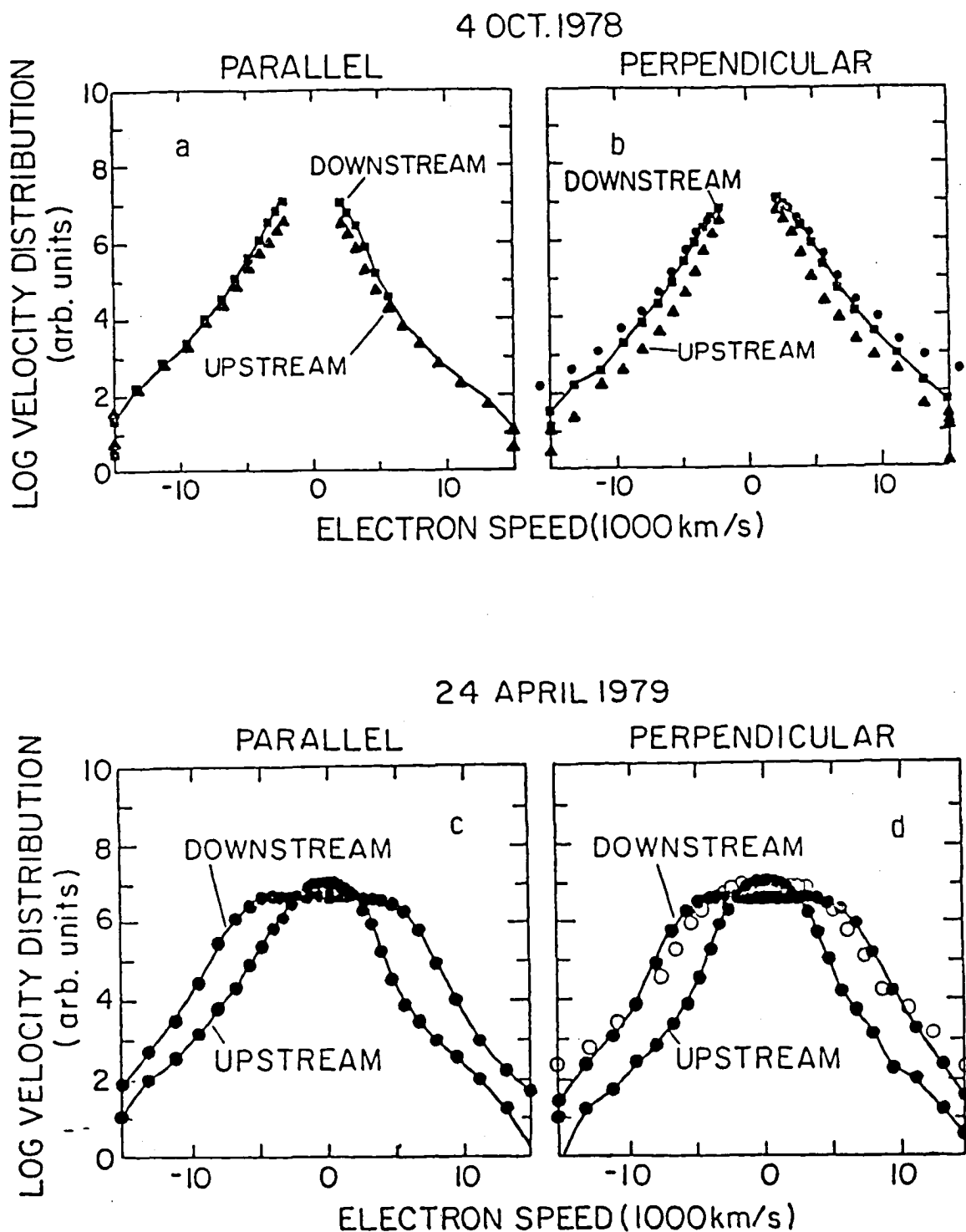


Figure 10-11. Cross sections of electron distributions across shock encounters of 4 October 1978 (upper panels) and 24 April 1979 (lower panels); "parallel" and "perpendicular" above the respective panels refer to cross sections parallel and perpendicular to the ambient magnetic fields.

parallel to  $\underline{B}$  might give rise to field-aligned ion acoustic waves (Thomsen et al., 1983b) and whistlers (Tokar et al., 1983). Finally, the presence of the reflected ions can give rise to, or modify the condition for lower-hybrid instability (Zhou et al., 1983a); kinetic cross-field streaming instability (Wu et al., 1983a); and the electron cyclotron drift instability (Zhou et al., 1983b). In the upstream region, field-aligned beams of electrons are observed to stream away from the bowshock back into the solar wind, potentially giving rise to electron plasma oscillations.

Evidence for the action of several of these instabilities has been found in light scattering observations of laboratory shocks and with plasma wave instrumentation on spacecraft. The Tarantula group carried out extensive studies of the microscopic turbulence in the resistive shock, using the scattering of ruby-laser light at small angles. From the spectrum of the scattered light it was deduced that the turbulence was due to current-driven ion waves having a Kadomtsev spectrum for  $k\lambda_D \sim 1$ . The magnitude of the fluctuations was consistent with the anomalous resistivity necessary to account for the shock width. When these measurements were extended to smaller  $k$  by using a  $\text{CO}_2$  laser (Craig et al., 1974), the fluctuations were found to be three orders of magnitude below that expected from an extrapolation of the Kadomtsev spectrum measured at  $k\lambda_D \sim 1$ .

Light scattering experiments in supercritical shocks were carried out by a group at Garching in a high- $\beta$  plasma (Keilhacker and Steuer, 1971). The turbulent fluctuations measured by scattering of  $\text{CO}_2$  laser light were almost identical to those observed in Tarantula, but because in this case  $T_e < T_i$ , the turbulence was attributed to electron-cyclotron waves.

Detailed measurements of the plasma wave spectrum in the Earth's bowshock (Fredricks et al., 1970; Rodriguez and Gurnett, 1975, 1976; Gurnett et al., 1979; Formisano and Torbert, 1982) have shown that a characteristic spectrum of electrostatic and electromagnetic plasma waves occurs in the bowshock, and consists of whistlers, ion acoustic-like waves, and electron plasma oscillations. Other wave modes that have not been conclusively identified also probably occur in the bowshock. Examples of plasma instabilities and their observational aspects may be found in Fredricks et al. (1970), Wu and Fredricks (1972), Greenstadt and Fredricks (1974), Lemons and Gary (1978), Wu et al. (1983b), (Thomsen et al. 1983) and in the preceding discussion.

The reviews by Greenstadt and Fredricks (1974), Wu (1982), and Wu et al. (1983b) list the following modes and their rest-frame frequencies: the Buneman mode with  $f \sim f_B = (m_e/m_i)^{1/3} f_{pe}$ ; ion acoustic waves with  $f \sim f_{pi}$ , lower hybrid drift waves with  $f \sim f_{LH} = (f_{ce} f_{ci})^{1/2}$ , and electron cyclotron drift waves with  $f \sim n f_{ce}$ , where  $f_{pe}$  and  $f_{pi}$  are the electron and ion plasma frequencies,  $f_{ce}$  and  $f_{ci}$  are the electron and ion cyclotron frequencies, and  $m_e$  and  $m_i$  the electron and ion masses. Observational difficulties are encountered in trying to identify these modes because the actual measurements include Doppler shifts, uncertain  $k$ -vector directions and multiple wave mode generation. However, shock plasma wave spectra do tend to span the range of frequencies listed above, which extend from  $f_{LH} \sim 5$  Hz to  $f_{pe} \sim 60$  kHz.

A series of spectra of electric field fluctuations in the shock transition region are shown in Figure 10-12 where spectrum sampling time is plotted on the oblique axis. Each spectrum is based on a 6-second average of the plasma wave power detected in the spectrum analyzer channels. The characteristic rest-frame frequencies listed above are indicated along the frequency axis of the 10th spectrum of Figure 10-12. Each shaded bar indicates the frequency range between average upstream and downstream values. The consistency of the spectral shape over many examples of bowshock crossings supports the following interpretation: the changing spectral shapes of the electric field fluctuations indicate the evolution of the bowshock plasma instabilities as a function of spatial position within the shock transition region.

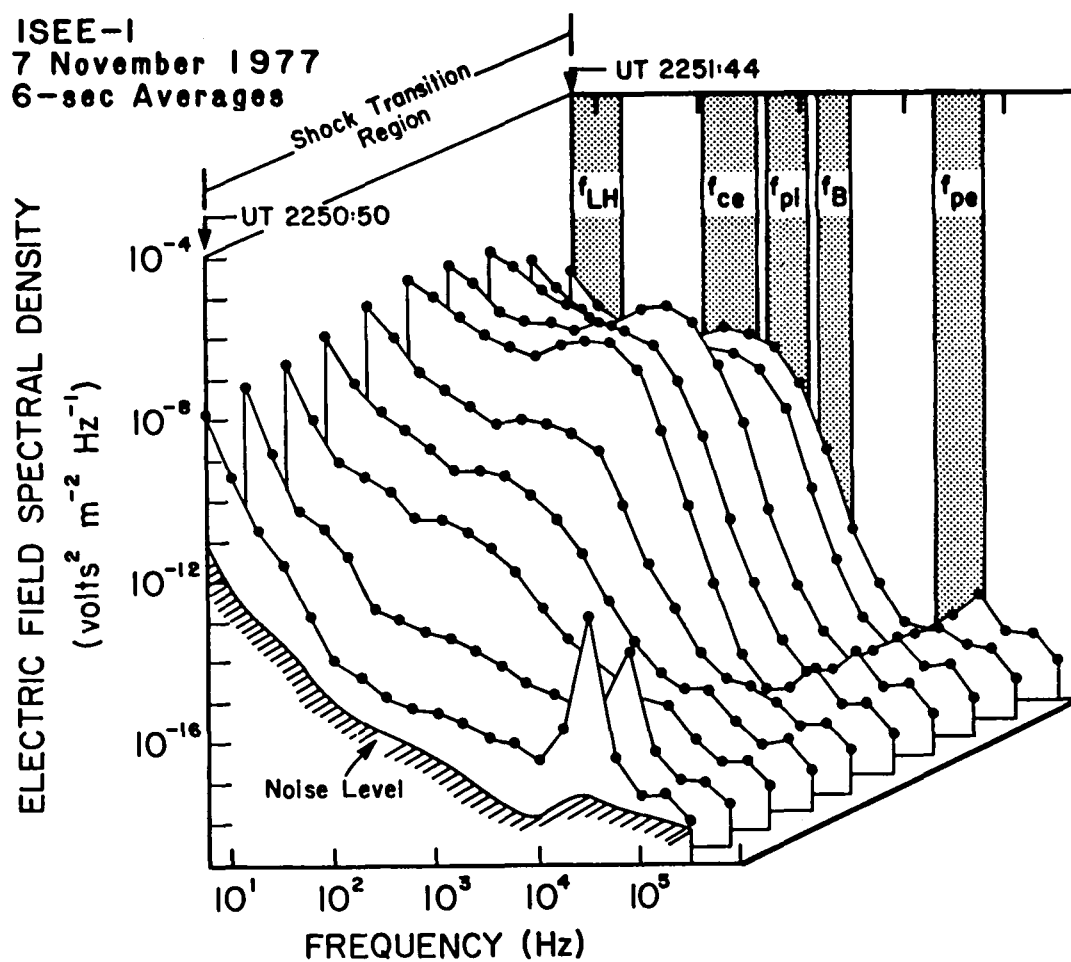


Figure 10-12. Time sequence of averaged plasma wave spectra through the shock crossing of 7 November 1977; shaded bars locate the (Doppler) ranges of various critical frequencies.

The intermediate frequency range  $100 \text{ Hz} < f < 10 \text{ kHz}$  in the bowshock plasma wave spectrum is characterized by a broad maximum which may be seen in Figure 10-12. At succeeding 6-second intervals (i.e., at succeeding locations further into the shock transition region) the power of electrostatic waves at frequencies below 10 kHz increases until the maximum broadband power occurs in the eighth spectrum. This spectrum has the enhanced electrostatic noise between 100 Hz and 10 kHz that is characteristic of bowshock spectra and has been discussed in detail by Rodriguez and Gurnett (1975). The intense electrostatic waves detected in the eighth spectrum of Figure 10-12 usually occur, as in this case, in the ramp of the magnetic field compression of the bowshock. These electrostatic waves are polarized primarily along the magnetic field direction, i.e.,  $\mathbf{k} \parallel \mathbf{B}$ , and have been identified as ion acoustic waves.

Low frequency waves,  $f < 100 \text{ Hz}$ , are also expected to be important in ion thermalization at the bowshock. Lower hybrid waves are considered as likely to be a significant wave mode (Lemons and Gary, 1978; Wu et al., 1983a). These waves may be generated by ion-ion streaming instabilities or by electron drift (current-driven) instabilities (Zhou et al., 1983a; Wu et al., 1983b). Lower hybrid waves can propagate perpendicularly or nearly perpendicularly to  $\mathbf{B}$  either in the shock plane or almost along the incoming plasma flow. Lower hybrid waves occur over a frequency range from 1-20 Hz in the plasma rest frame and over a larger range if Doppler shifting is a significant factor, perhaps to frequencies greater than 100 Hz. The low frequencies in Figure 10-12 clearly show an enhancement in the shock transition; however, frequency resolution is limited. Whistler waves have also been suggested as a possible wave mode, the identity of which rests primarily on the observed cutoff of low frequency magnetic field fluctuations at  $f \sim 100\text{-}300 \text{ Hz}$  (Rodriguez and Gurnett, 1975). Detailed quantitative observations of bowshock plasma waves at frequencies  $f < 20 \text{ Hz}$  have only recently begun (Formisano and Torbert, 1982; Vaisberg et al., 1983), and much work remains to be done. The low frequency part of the spectrum is potentially an area of significant contribution to the understanding of collisionless shock structure.

The high frequency regime in Figure 10-12 ( $f > 10 \text{ kHz}$ ) is associated with electron plasma oscillations which commonly occur in the foot of the bowshock and also at appreciable distances upstream. These oscillations are thought to be caused by the backstreaming energetic electrons (Fredricks et al., 1970; Wu et al., 1981). The frequency of electron plasma oscillations is proportional to the total plasma density. The peak at about  $f_{pe} \sim 30 \text{ kHz}$  occurring in the first two spectra in Figure 10-12 represents these electron plasma oscillations. Further into the shock transition, the peak decreases and moves to a higher frequency ( $f_{pe} \sim 60 \text{ kHz}$ ) since the plasma density is increasing. Doppler shifts are generally insignificant for electron plasma oscillation because of the high rest frame frequency.

#### IV. QUASI-PERPENDICULAR SUBCRITICAL PROCESSES

In ideal subcritical shocks, by definition, the dissipation necessary to limit shock steepening should arise entirely from wave dispersion and/or in anomalous resistivity, in contrast to supercritical shocks where an additional dissipation, provided by gyrating reflected ions, combines with resistivity to form a finite profile. Subcritical shock waves have received a great deal of attention, especially those of both low  $\beta$  and low Mach number, which have been designated laminar shocks. Laminar shocks have a well developed theory (see for example Biskamp, 1973 and Galeev, 1976) and they have been extensively studied in laboratory plasmas (see reviews by Robson, 1969; Eiselevich, 1982). The picture which emerges from this work is one in which shock structure reflects a competition between resistive and dispersive effects.

By competition we mean the following: a nonlinear wave will tend to steepen as it propagates because phase velocity increases with amplitude, causing larger amplitude parts of the wave to overtake smaller amplitude parts. This nonlinear steepening can be arrested by two possible processes: (1) If the wave is dispersive, i.e., if the phase velocity depends on frequency, the higher frequency components will tend to run ahead of the wave or lag behind it, forming a leading or trailing wave train, respectively, rather than contributing to a steep front; (2) Some plasma instability may give rise to anomalous resistivity within the shock so that the shock current (responsible for the increase in magnetic field strength constituting the front) cannot increase indefinitely, thereby limiting the shock jump, and thickness, to a finite value.

According to this picture, then, in perpendicular and nearly perpendicular shocks the magnetic field gradient is relatively steep and the associated current exceeds the threshold for the excitation of electrostatic waves. The anomalous resistivity generated by these waves then limits the current, and the resistivity is thus responsible for limiting the shock steepness. Current driven ion acoustic turbulence (Galeev, 1976); electron beam driven ion acoustic waves (Thomsen et al., 1983c) and lower hybrid drift modes (Lemons and Gary, 1978) have all been suggested as sources of the electrostatic turbulence in perpendicular laminar shocks. In more oblique shocks, however, field gradients are not steep enough to support sufficient currents to drive strong electrostatic waves, and shock steepening and structure are controlled by the dispersive character of the plasma. In this case, as first discussed by Sagdeev (1966), the shock structure is essentially that of a large amplitude damped whistler wave.

Recent studies of a set of oblique laminar bowshocks based on the ISEE-1 and -2 dual spacecraft data set have confirmed many elements of this picture (Mellott and Greenstadt, 1984). These measurements have documented the whistler nature of the oblique shocks and shown that they are characterized by resistive length scales several times greater than the dispersive scale of the plasma. The increase in damping with the upstream  $T_e/T_i$  in this study is suggestive of the importance of ion acoustic turbulence in these shocks, but the source of the anomalous resistivity in these shocks is not clearly established.

An earlier study of this same set of laminar bowshocks (Russell et al., 1982) revealed that although electron heating was comparable to what is commonly observed at the supercritical bowshock, the ion heating was substantially reduced, as would be expected in the absence of reflected, gyrating ions. A recent study of subcritical interplanetary shocks (Feldman et al., 1983b) suggests that the shapes of observed downstream electron distributions can for the most part be accounted for by conservation of magnetic moment alone. Any acceleration in the parallel direction associated with the shock potential drop is much less significant than for supercritical shocks.

Of the sources of particle disequilibrium described above for supercritical quasi-perpendicular shocks, only two seem to be applicable at subcritical shocks, namely the increase in perpendicular electron temperature and the cross-field current associated with the increase in the field strength. However, the nonlinear effects of microscopic instabilities are not apparent in the electron distribution function. Whichever instability is responsible for limiting the cross-field current seems to do little electron heating in the process.

## V. QUASI-PARALLEL PROCESSES

Mathematically and conceptually, studies of shocks in a magnetized plasma naturally lean toward defining two ideal extremes: perpendicular and parallel. In the first case, the magnetic field is a

fundamental organizer of wave propagation and particle distributions; in the second case, the field is essentially invisible to the plasma flow. Any case in between, in the absence of an a priori reason to single out a particular direction, is simply oblique. Even if we grant special status to the perpendicular and parallel extremes, however, we intuitively expect some gradation of properties depending on the degree of obliquity. But if such a gradation exists in the usual supercritical regime of space, it is subtle in comparison with what spacecraft observations have revealed to be the major dichotomy between quasi-perpendicular and quasi-parallel shock structure in nature.

## TERMINOLOGY

A description of the source of the term quasi-parallel will be helpful and will convey the context in which we are almost always obliged to investigate the oblique phase of the Earth's bowshock.

Communication of a shock with its own upstream region under parallel conditions, with implied nonuniformity of the bowshock, was originally attributed to the dispersive character of wave propagation in the solar wind plasma. Parallel traveling waves with group velocity components greater than the projection of the solar wind's velocity on the IMF would presumably be capable of propagating upstream. Such dispersion would remove energy from the shock in the form of waves and alter its structure. Based on this reasoning, Kellogg (1964) correctly predicted the asymmetry of the bowshock in the simple steady state geometry of the average IMF. However, there was no theoretical prediction of just how the profile might depend locally on the deviation from exact parallelism, and no prediction of the profile that was actually encountered in space.

It was clear, to the first instruments with the appropriate resolution, that magnetic profiles of the bowshock for low  $\theta_{BN}$  were grossly different from profiles for large  $\theta_{BN}$ , that the more "parallel" profiles were very thick magnetically, and that the bowshock was indeed asymmetric and, by inference, locally nonuniform (Greenstadt et al., 1970). Additional correlations were found: ULF waves appeared adjacent to the "shock" when, and only when, the latter exhibited a turbulent, highly disturbed profile characterized as the "thick pulsation shock" (Greenstadt et al., 1970, Part 2); in addition, the waves, whether near or far from the nominal shock, were recorded only behind a tangent line to the shock at whose point of contact  $\theta_{BN}$  was consistently less than  $\sim 51^\circ$  (Greenstadt et al., 1970, Part 2; Diodato et al., 1976); finally, no obvious distinction between "pulsation" profiles was apparent for different local, nominal  $\theta_{BN}$  values. It followed that the bowshock generally had a large section of "pulsation" profile, including the local points of strictly "parallel" geometry, extending toward the locally "perpendicular" geometry as far as, but only as far as, the points where upstream ULF waves were to be found. This was the basis of the designation "quasi-parallel" for the shock when  $\theta_{BN} \lesssim 50^\circ$ . It must still be thought of as an empirical designation applicable to portions of natural shocks such as the bowshock, where large amplitude ULF waves of mixed periods and irregular shapes adjoin, and sometimes alternate with, ULF wavetrains of smaller amplitude.

## ION PARTICIPATION

Well before the geometric division of structures was noted, backstreaming ions with energies greater than those of the solar wind had been measured in satellite data (Asbridge et al., 1968). A connection between escaping particles and waves was subsequently discerned in space, forced by the observation of waves of decasecond (ULF) periods that could not possibly have propagated upstream from the bowshock, but that were invariably seen when the IMF intercepted to the shock (Fairfield, 1969). These waves were being carried downstream by the solar wind but had to have been generated by ions coming from the direction of the shock. It was also recognized in the laboratory, as discussed in Section III above, that "oblique" shocks had upstream precursors and that supercritical shocks were distinguished by the presence of reflected ions,



some of which could be presumed to escape the shock and find their way upstream. Certainly when ions do not escape, they have an appreciable effect on the shock profile (e.g., the foot, the overshoot), and it may be assumed their escape would alter that profile, much as the dispersion of wave components along the ambient field should alter the profile of the parallel shock.

One and two-dimensional numerical simulations of uniform parallel and approximately parallel shocks have shown both the escape of ions back into the upstream plasma and the development of magnetic profiles featuring whistler mode waves of low frequency and large amplitudes. These waves have irregular forms caused by the admixture of wave components of higher frequencies (Biskamp and Welter, 1972b; Forslund et al., 1983; Quest et al., 1983; Kan and Swift, 1983; Leroy and Winske, 1983). Thus, we have reason to expect a distinguishable magnetic shock structure associated directly or indirectly with ion escape and upstream waves, and including within its span the local conditions formally linked to parallel geometry; in short, a quasi-parallel structure.

## MINGLED PROCESSES

Theoretical treatments of oblique shocks were summarized in Chapters 4, 5, and 9 of Tidman and Krall (1971), with most of the emphasis on laminar shocks and on whistler mode structure. Subsequent papers, for example, that of Cipolla et al. (1977), have explored further the effects of finite ion temperatures and counterstreaming ion instabilities in producing large amplitude wave structures in realistic plasmas. The firehose instability has played a role in parallel shock theory (Kennel and Sagdeev, 1967; Auer and Volk, 1973) with mixed success when tested with space observations (Greenstadt et al., 1977; no apparent firehose participation) and simulation (Kan and Swift, 1983; definite firehose participation). All these results, of course, suggest an intrinsic parallel, oblique, or quasi-parallel shock structure.

As documented in other sections, one of the most consistent and now thoroughly established features of supercritical shocks is the reflection of a fraction of the incoming ions to create a secondary distribution of particles which, under the appropriate circumstances, escape the shock into the upstream region. The consequences of this phenomenon in nature, are elaborated in Section VI, below. The essential implication is: Some of what the parallel side of a curved shock sees is the result of what the perpendicular side creates. Hence, it is difficult to document observations of the intrinsic structure of the quasi-parallel portion of a nonuniform shock.

One way of attempting to separate the intrinsic from the confounded components of the oblique shock is to look for shocks of larger radius of curvature than that of the Earth, in hope of approximating the ideal plane shock. The most accessible of these are the interplanetary shocks, of which many have been recorded by spacecraft instruments in the solar wind. It appears at present that quasi-parallel interplanetary shocks do exhibit intrinsic foreshocks, but that there are distinguishable differences between these foreshocks and the one commonly associated with the Earth's bowshock (Kennel et al., 1982; Tsurutani et al., 1983; Russell et al., 1983). The most notable difference is the relative absence of compressional waves ahead of interplanetary events. A comprehensive separation of effects remains to be achieved.

## PHENOMENOLOGY

The complexity of quasi-parallel structure and the practical obstacles to describing and measuring them have induced investigators to postpone their study repeatedly. There is not therefore a large body of literature on the subject. We summarize what there is at present, as fresh studies begin in earnest.

Because orbiting spacecraft accumulate hundreds of shock crossings in their observational lifetimes (Figure 10-1), the bulk of all observations of natural shocks necessarily involve planetary bowshocks, in particular those of Venus and Earth, Venus' shock has hardly been examined structurally at this time, and we are just beginning to study in detail the oblique side of the Earth's bowshock. We review briefly the record on the latter.

We note first some conditions:

- There has been no general documentation of a dependence of quasi-parallel properties on other plasma parameters, that is, on  $\beta$ , Mach number,  $T_e/T_i$ , etc.
- Most of what has been described to date has been the shock profile as viewed by magnetometers. This has been largely because of the slow sampling cycles of plasma instruments relative to typical wave periods in the magnetic structure.
- The thickness of the pulsation structure and the effect of the foreshock on the solar wind have made estimation of true upstream parameters inaccurate at best and possibly misleading in many cases.
- The irregular and unreproducible magnetic profiles of individual shock encounters have made estimates of shock motion and local orientations unreliable, even when the pair of satellites ISEE-1, -2 have sampled the same shock crossing.

Subject to these conditions, we have established the following:

First order magnetic profiles of subcritical shocks, all interplanetary, show definable shock jumps with clearly distinguished upstream and downstream field levels (Tsurutani et al., 1983). This property contrasts with the appearance of many bowshock examples, and is probably attributable in part to the high speed with which interplanetary shocks pass the spacecraft.

Supercritical shocks of moderate  $\beta$  are characterized by (1) irregular large amplitude magnetic pulsations, sometimes in bursts, often separated by intervals of smaller amplitude upstreamlike waves; (2) thickness of  $\gtrsim 2RE$ ; (3) large amplitude quasi-periodic transverse magnetic wave components; (4) solar wind of elevated temperature, enhanced density, and distinct distribution with skewed high-energy tails and irregular low-energy envelopes; (6) inversely related antisolar-directed plasma flux and field magnitudes; (7) electric and magnetic ELF wave noise comparable to that associated with laminar shocks; (8) interpulsation regions of upstream magnetic magnitude and wave structure but noisy, deflected, and partially thermalized plasma flow; (9) macrostructure following the outlines of an oblique whistler shock, modified by additional irregularity and complexity (Greenstadt et al., 1977).

The large amplitude pulsations have a correlation length of several hundred kilometers, which may be related to the gyroradii of ions of the escaping secondary distribution (Greenstadt et al., 1982).

## VI. ION ACCELERATION

### THE FORESHOCK

We describe the observed particles and waves of Earth's foreshock, noting that similar constituents are observed upstream of interplanetary traveling shocks. However, foreshock and (quasi-parallel) shock structures depend significantly on geometrical effects. A relatively small shock radius of curvature relative to the scale over which foreshock processes take place, as in the case of Earth's bowshock, leads to complexities not expected in interplanetary shocks more closely approximating plane surfaces over the same scale or in astrophysical shocks.

The geometry of the terrestrial foreshock is indicated in Figures 10-13 and 10-15. That interplanetary magnetic field line which is just tangent to the shock forms the limiting upstream foreshock particle boundary: No particle with source in, or downstream of, the shock can propagate upstream of this boundary, although we may imagine a continuous set of sub-boundaries downstream of the foreshock boundary corresponding to particles with different energies moving out from the shock. Each sub-boundary is determined by the sum of field-aligned particle motions and the  $E \times B$  drift. In the energy ranges encountered in the terrestrial foreshock, drift effects are more substantial in the case of ion than of electron motion, and the ion foreshock boundary is found significantly downstream of the electron boundary.

For instruments moving downstream into the foreshock, the first particles encountered are beams of energetic ( $E > 16$  KeV) electrons. Sub-boundaries of successively lower energy electrons are encountered in a pattern consistent with the sweeping back of electron trajectories by the solar wind electric field (Anderson et al., 1979). At still lower energies, another typical foreshock electron signature is that of an electron heat flux directed back towards the Sun, in contrast to the case in the undisturbed solar wind where the heat flux flows away from the Sun (Ogilvie et al., 1971; Feldman et al., 1982). The development of anisotropies in electron distributions as a function of location within the foreshock are discussed in detail by Feldman et al. (1983a).

Foreshock ion distributions also change in both energy and character downstream from the ion foreshock boundary (Figure 10-4). First encountered are nearly field-aligned beams of ions with typical energies of several KeV. Further downstream, ion distributions are spread in energy and pitch angle, forming kidney-bean shaped ("intermediate") distributions and, finally, deep in the foreshock, the ions form nearly isotropic ("diffuse") populations (Gosling et al., 1978; Greenstadt et al., 1980a). The spectrum of the latter population typically extends up to several hundred keV (e.g., Scholer et al., 1979; Ipavich et al., 1981a). In addition to the above populations, "gyrating" ions or phase-bunched beams are often observed deep within the foreshock (e.g., Gurgiolo et al., 1981; Eastman et al., 1981).

A variety of waves are found in association with the foreshock particle distributions (Paschmann et al., 1979; Anderson et al., 1981; Hoppe et al., 1981). Some waves appear to have analogs in the undisturbed solar wind and some are unique to the foreshock. The highest frequency emissions are impulsive narrow band electrostatic signals at the local electron plasma frequency. There are most intense in the vicinity of the foreshock boundary and are identified as electron plasma oscillations usually presumed to result from a two stream instability between backstreaming electron beams and the incoming solar wind. There are also suggestions that there may also be enhanced generation at particle flux gradients. A second more broad band electrostatic signal with frequencies in the range  $0.5 \text{ kHz} < f < 5 \text{ kHz}$  has been identified as Doppler-shifted ion acoustic waves occurring in conjunction with foreshock ion distributions of all classes. These may theoretically be generated by either electron heat fluxes or ion beams (Lemons et al., 1979), but so far observed foreshock particle distributions appear to be stable to

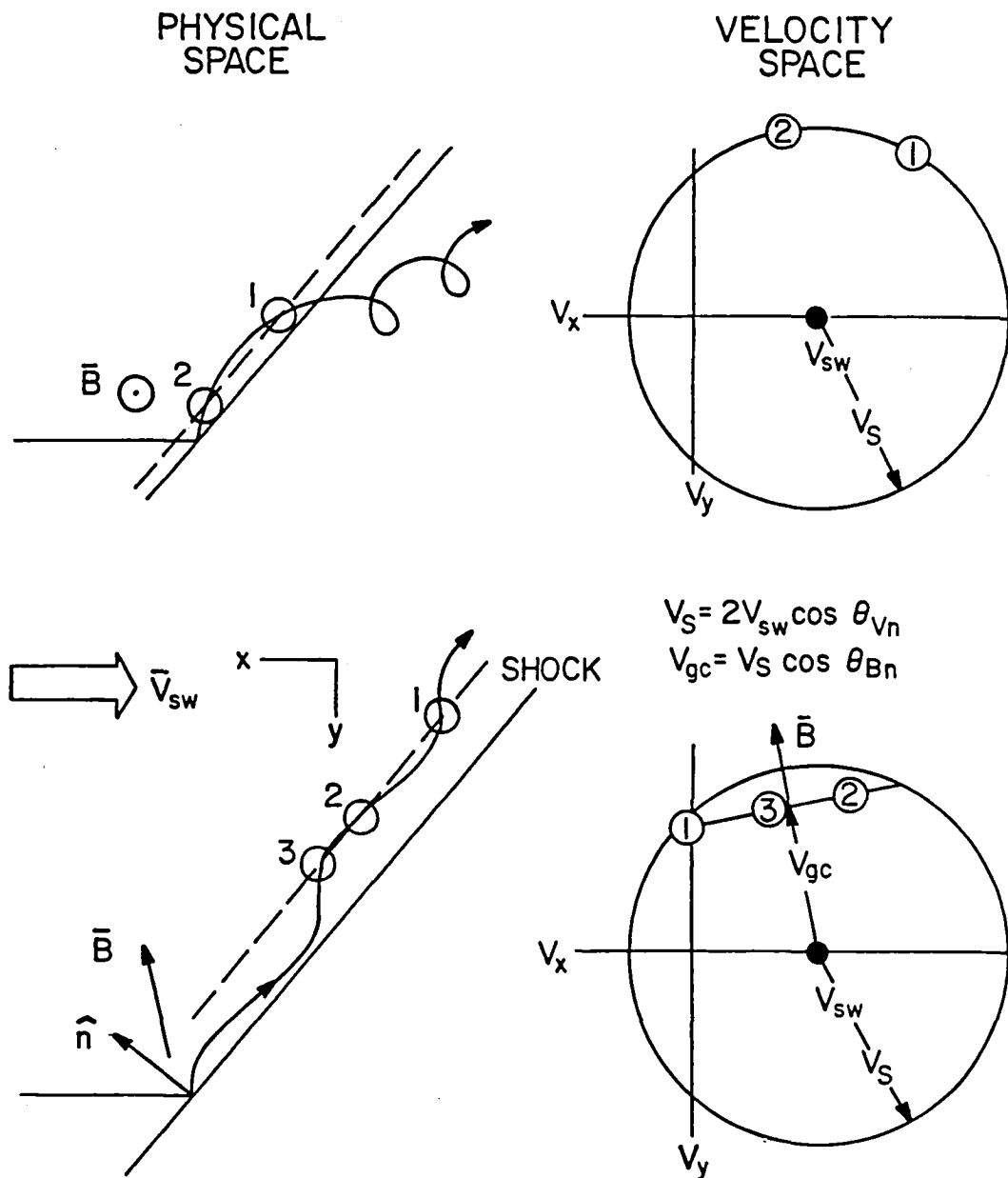


Figure 10-13. (Left) Trajectories, in physical space, of specularly reflected ions for different orientations of interplanetary field  $B$  with respect to the shock; (Right) velocity space diagrams of positions where reflected ions should be expected to appear in two-dimensional contour distributions recorded at the locations indicated by corresponding numbers at left.

both mechanisms, and further work on identifying the wave sources is needed. The third foreshock plasma wave emission, a low frequency ( $f < 200$  Hz) electromagnetic wave, has been identified on the basis of electromagnetic field measurements as whistler mode emission. Its source is not yet established.

At still lower frequencies are found a variety of electromagnetic waves. In the vicinity of the ion foreshock boundary, there appear relatively small amplitude ( $\delta B/B \sim 0.1$ ), obliquely propagating waves with frequencies  $\sim 1$  Hz. These waves appear to be driven by features of the foreshock electron distributions (Feldman et al., 1983a; Sentman et al., 1983). Further downstream, in the region populated by intermediate and diffuse ions, appear large amplitude ( $\delta B/B \sim 1$ ) low frequency ( $f \sim 0.3$  Hz) upstream waves (Greenstadt et al., 1968; Fairfield, 1969). Hoppe and Russell (1983) have shown that while the majority of these signals are magnetosonic mode waves with rest-frame frequencies  $\sim 0.1 \Omega_p$ , they are occasionally found propagating in the Alfvén mode with similar rest-frame frequencies and wavelengths. The generation of the magnetosonic form of these waves can be explained by cyclotron resonant instability driven by reflected ion beams (Barnes, 1970; Gary et al., 1981) but the observation of the Alfvén mode suggests wave generation by the more isotropic diffuse distributions as well (Sentman et al., 1981).

The most detailed studies of foreshock phenomena have been conducted on data from the terrestrial and Jovian bowshocks (Hoppe and Russell, 1982; Zwickl et al., 1980, 1981; Smith et al., 1983), but similar effects are now being documented upstream of interplanetary shocks. Tsurutani et al. (1983) and Russell et al. (1983) for instance, have shown that low frequency waves similar to those studied in the terrestrial foreshock appear upstream of some interplanetary shocks. Gosling et al. (1983), on the other hand, have reported that the suprathermal ions observed upstream of interplanetary shocks, while sharing some of the features seen in the terrestrial case, also exhibit some differences in character. One interesting feature of the regions upstream of interplanetary shocks is the large length scales of some shock-associated effects (Kennel et al., 1982).

## PROTON INJECTION INTO THE UPSTREAM FLOW

A number of suggestions have been made concerning the possible origins of the various ion distributions observed in the foreshock. However, the most promising argument at present is that all of these distributions arise from the initial nearly specular reflection of a small fraction of the solar wind ions incident on the bowshock (see Section III.)

The trajectory of an ion following specular reflection depends upon  $\theta_{Bn}$ . For  $\theta_{Bn} > 45^\circ$  (quasi-perpendicular shock) the guiding center motion is directed downstream, whereas for  $\theta_{Bn}$  (quasi-parallel shock) the guiding center motion is directed upstream (Gosling et al., 1982). The left-hand portion of Figure 10-13 shows trajectories of specularly reflected ions for  $\theta_{Bn} = 90^\circ$  (upper panel) and for  $\theta_{Bn} = 37^\circ$  (lower panel).

In the upstream rest frame, specularly reflected ions initially have a speed  $2 V_{sw} \cos \theta_{vn}$  directed parallel to the shock normal, where  $\theta_{vn}$  is the angle between the incident upstream velocity vector and the local shock normal. Because energy is conserved in the upstream rest frame, the ion's subsequent trajectory in velocity space is on a sphere of radius  $V_s = 2 V_{sw} \cos \theta_{vn}$ , whose intersection with the xy plane is shown as a large circle in the velocity space diagrams at the right in Figure 10-13. The trajectory can be decomposed into a guiding center motion,  $v_{gc}$ , plus a gyration motion,  $v_g$ . In the upstream solar wind frame (Gosling et al., 1982)

$$v_{gc} = 2 V_{sw} \cos \theta_{vn} \cos \theta_{Bn}$$

and

$$v_g = 2 V_{sw} \cos \theta_{vn} \sin \theta_{Bn}$$

Note that for a perpendicular shock,  $\theta_{Bn} = 90^\circ$ , (upper panel, Figure 10-13) the motion is entirely gyrational in the solar wind frame, whereas for a parallel shock  $\theta_{Bn} = 0^\circ$ , the reflected ions have zero gyromotion. The small numbered circles in the right-hand portion of Figure 10-13 are the positions in 2-dimensional velocity space where specularly reflected ions would be observed at the upstream distances indicated by the dashed lines in the left-hand side of the figure. Thus specular reflection produces gyrating ions for widely different values of  $\theta_{Bn}$  whose phases depend upon distance from the shock and which are in reasonable quantitative agreement with observations (Paschmann et al., 1982; Gosling et al., 1982; Sckopke et al., 1983).

For  $\theta_{Bn} > 45^\circ$ , the normal velocity component (in the shock frame) of a specularly reflected ion returning to the shock exceeds  $V_{sw} \cos \theta_{Bn}$ , so such ions should continue to gyrate downstream (Schwartz et al., 1983) as described in Section III. Calculations indicate that the gyrating ion population will excite electromagnetic ion cyclotron waves which, in turn, pitch angle scatter the gyrating ions within a few gyroperiods; as a result, some of the specularly reflected ions in the downstream region acquire large speeds parallel to  $\underline{B}$  which can be sufficient to carry them back to the shock to escape into the upstream region (Tanaka et al., 1983).

The situation for the bowshock is illustrated schematically in Figure 10-14 (from Thomsen et al., 1983b). A source region, represented as a surface of scattered, gyrating ions, exists several ion gyroradii downstream from the shock (dashed line). The lines drawn tangent to the source surface and labelled with various values of the parameter  $\alpha$  represent the most sunward trajectories of shell ions leaving the source surface with pitch angle  $\alpha$  (measured in the downstream plasma rest frame; see inset 4 on the right-hand side of the figure). For this geometry, no particles from the source surface can escape duskward of point A on the shock. As noted by Tanaka et al., (1983) this provides a natural explanation of the fact that the upstream field-aligned beams are not observed for  $\theta_{Bn} > 75^\circ$  (e.g., Paschmann et al., 1981; Bonifazi and Moreno, 1981a).

The 2-dimensional phase space diagrams at the right of Figure 10-14 indicate schematically the ion distributions that would be observed along the nominal spacecraft trajectory 1-5 shown in the left-hand portion of the figure. For simplicity the upstream solar wind beam and the directly transmitted downstream solar wind beam are indicated by small dots. Downstream of the source surface (position 5), the entire shell of scattered, gyrating ions is observed. Upstream of the source surface, but downstream of the shock (position 4) a broad, field-aligned beam of particles is observed moving toward the shock. Immediately upstream of the shock (position 3) an escaping beam of particles is observed with  $T_\perp > T_\parallel$ . This escaping beam has a slightly higher energy than the beam at position 4 and is shifted slightly off the field direction because of an acceleration through the shock potential parallel to the shock normal as the ions cross the shock (Schwartz et al., 1983; Thomsen et al., 1983b). At position 2 the beam has a small  $T_\perp$  but is still shifted slightly off the field direction, and at position 1 the spacecraft is outside the region accessible to ions from the source surface.

The above sequence is consistent with the observed evolution of nearly field-aligned ion beam distributions across the shock (Thomsen et al., 1983b), and the upstream energies predicted by this sequence are consistent with observed energies of most upstream "field-aligned" beams (Thomsen et al., 1983a). However, the above leakage model of ions which are initially specularly reflected at the shock cannot explain the highest energy beams observed upstream (Tanaka et al., 1983; Thomsen et al., 1983a). Multiply reflected ions are possibly the source of these higher energy beams (e.g., Greenstadt et al., 1980a; Thomsen et al., 1983a; Leroy and Winske, 1983).

Examination of Figure 10-14 suggests that in the quasi-parallel regime the leakage back across the shock of ions which are originally specularly reflected at the quasi-perpendicular shock should produce upstream ion distributions which resemble the "intermediate" ion distributions shown in

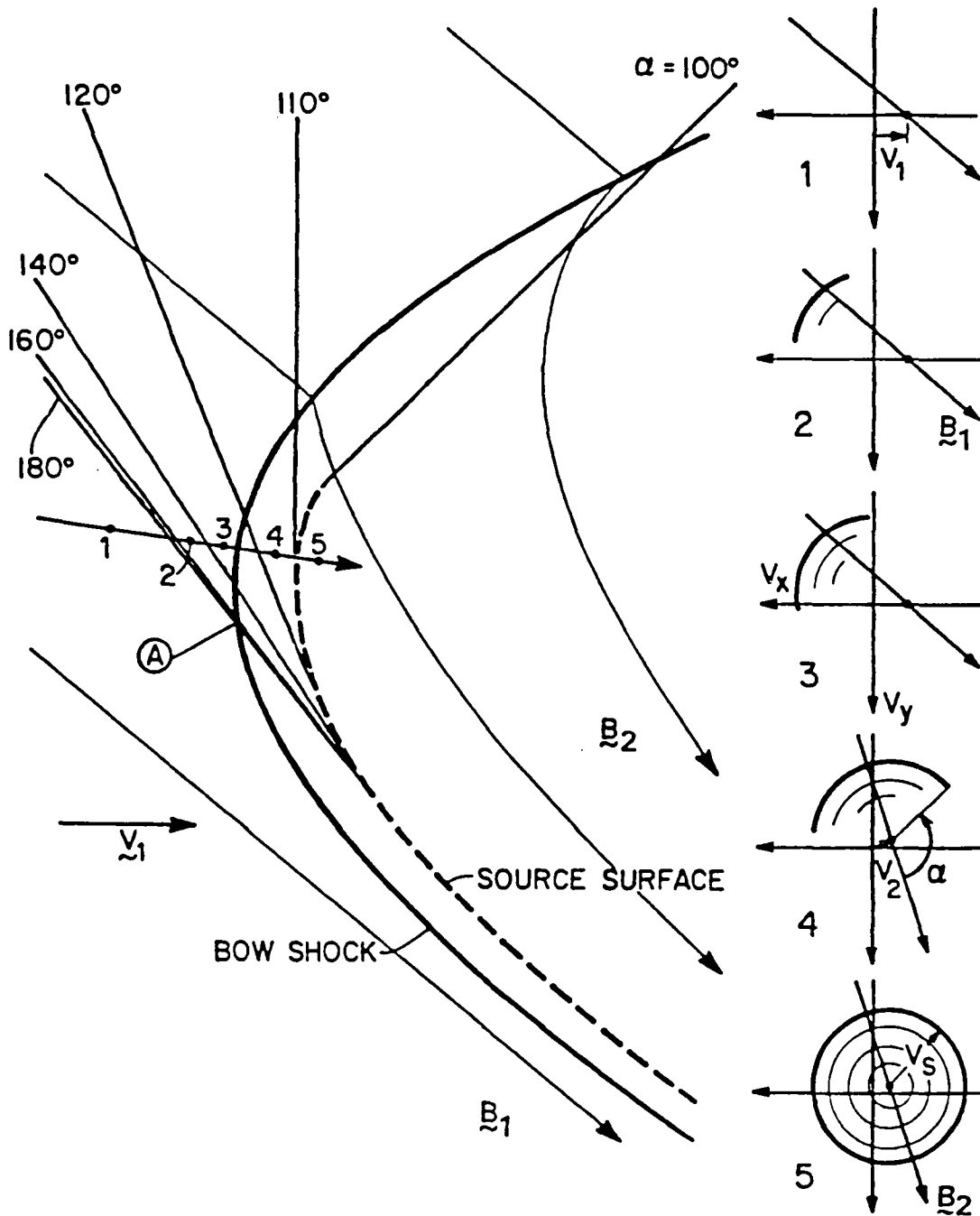


Figure 10-14. (Left) Trajectories of  $\alpha^+$  ions of various pitch angles returned to the solar wind from the source surface or reflected ions heated and scattered behind the quasi-perpendicular section of the bowshock; (Right) velocity space diagrams of expected contour locations of return ions in distributions which might be obtained at the corresponding numbered locations at left.

Figure 10-4. A similar explanation of intermediate type distributions has been given by Edmiston et al. (1982). Alternatively, it has been suggested (e.g., Bame et al., 1980; Bonifazi and Moreno, 1981b; Paschmann et al., 1981; Gary et al., 1981) that intermediate ion distributions are produced by the pitch angle scattering of field-aligned beams in the upstream region as we shall discuss in the next section. Quite possibly all of the above processes contribute at one time or another to the production of the intermediate ion distributions.

## ION ACCELERATION AND HYDROMAGNETIC WAVE EXCITATION

Figure 10-15 depicts Earth, the bowshock, and a plane containing the solar wind velocity vector,  $\underline{V}_{SW}$ , and the average magnetic field,  $\underline{B}$  (for convenience we have chosen that plane which passes through the Earth as well). The angle  $\theta_{Vn}$  is the angle between  $\underline{V}_{SW}$  and  $\underline{n}$ ;  $\underline{n}$  denotes a shock normal which makes an angle  $\theta_{Bn}$  with  $\underline{B}$ . The figure is approximately to scale; the nose of the bowshock is typically  $14 R_E$  from Earth.

When a magnetic field line convecting with the solar wind first contacts the bowshock at P, the escaping field-aligned ions we have described first have access in principle to the upstream region, although with typical energies of 4-5 keV a substantial number can escape only at P' where  $\theta_{Bn} \sim 75^\circ$ . These escaping ions determine the foreshock boundary, F, the guiding center trajectory of a typical unscattered ion. The streaming ions interact with interplanetary waves in cyclotron resonance, damping the majority which propagate away from the Sun relative to the solar wind (and which are left-hand polarized in the frame of the solar wind) and exciting the minority which propagate in the opposite direction (right-hand polarized). There should result initially a decrease in wave intensity (although this has not been observed) and a reduction in ion scattering, thus creating a sheath within the foreshock boundary characterized by reduced wave intensity and "free-streaming" ion distributions, R, with a characteristic velocity distribution as shown at the top of the figure. Following a given field line as it convects with the solar wind, the unstable predominantly right-hand polarized wave minority eventually grows to sufficient intensity to begin to scatter the ions in pitch angle back toward the shock front. There results a sheath of nearly monochromatic (due to the relatively narrow ion velocity distribution) right-hand polarized waves as shown and partially scattered "intermediate", I, ion distributions, a sample of which is depicted at the top of the figure.

Eventually, the enhanced wave intensity scatters a substantial number of ions back toward the shock front creating a more isotropic ion distribution. These scattered ions excite predominantly left-hand polarized waves propagating away from the bowshock relative to the solar wind. There results a sheath of less strongly polarized waves and "diffuse" ion distributions, D, confined to lower energies (<30 keV) as shown in the figure. Finally field line connection time is sufficiently large near the bowshock that some ions have been scattered across the shock ("subshock") several times between the foreshock and the downstream turbulence. In the process the particles are accelerated by shock compression and drift parallel to the average electric field. Viewed in a frame moving along the shock front in which the electric field vanishes the particles gain energy by being scattered between converging "walls", the defining characteristic of first-order Fermi acceleration. These ions are accelerated up to ~120 keV and form the "diffuse",  $\bar{D}$ , ion distributions as shown. The associated ULF waves have evolved (presumably via nonlinear interactions) to more complex structures exhibiting compression, wavevectors oblique to  $\underline{B}$ , both polarizations, and the presence of "shocklets". This scenario is supported qualitatively by the foreshock observations we have described of the ions and ULF waves and their correlations.

A quasi-equilibrium can result in sheath D in which ion acceleration is balanced by ion loss due to upstream escape along the field lines, or lateral escape onto field lines disconnected from the nose of the bowshock where compression is substantial by either spatial diffusion perpendicular to  $\underline{B}$ , field line convection, or ion drift in the inhomogeneous average magnetic field,  $\underline{B}$ , at



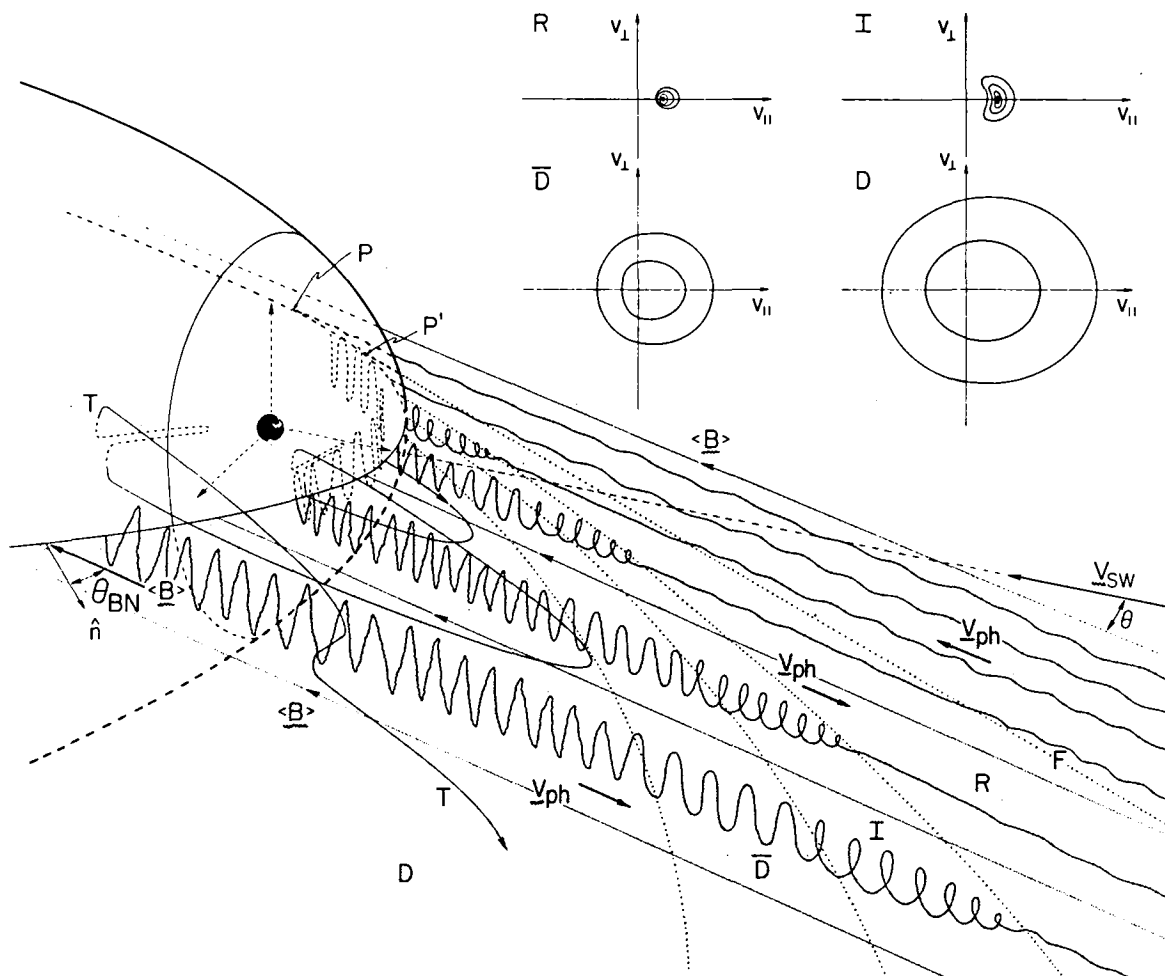


Figure 10-15. Conceptual diagram of foreshock ion Fermi-acceleration in one plane; velocity space diagrams inserted at upper right indicate the evolution of return ion distributions at the corresponding lettered locations in the lower picture.

the shock front. A typical ion guiding center trajectory (T) is shown in the figure including injection from the downstream plasma, four shock encounters, reflection from the downstream turbulence, and drift at the shock front. The waves are simply convected with the solar wind and compressed through the shock front, contributing substantially to the downstream turbulence as indicated in the figure.

A self-consistent theory of ion acceleration and ULF wave excitation in the quasi-equilibrium has been developed neglecting nonlinear wave-wave and wave-particle interactions and assuming given ion injection rates which in principle are provided by the physics presented in the last subsection (Lee, 1982). The theory is in quantitative agreement with the ISEE observations of the ion omnidirectional distributions as functions of energy and distance upstream of the bowshock and of the ULF wave power spectral density as a function of frequency and distance upstream.

Similar configurations of ULF waves (Tsurutani et al., 1983) and energetic ions are observed in association with interplanetary traveling shocks. The ion enhancements are known as energetic storm particles (ESP) and post-shock ion enhancements (Scholer et al., 1983; Gloeckler et al., 1974; Gosling et al., 1981). Since the connection time of magnetic field lines to the shock front is generally longer than the characteristic acceleration times, however, the "reflected" and "intermediate" ion distributions are not observed. Rather the observed ion distributions are often more nearly isotropic in the solar wind or shock frames and correspond to the "diffuse" ion distributions at the bowshock. A self-consistent theory for ion acceleration and ULF wave excitation at interplanetary traveling shocks has been developed for the nearly isotropic events (Lee, 1983) which is in very good agreement with the characteristic behavior of the observed ion omnidirectional distributions and wave power spectral densities. Detailed application of the theory to the large supercritical 11, 12 November 1978 event yields excellent agreement with the observed scalelength for ion intensity decrease upstream of the shock, the power law spectral index, the wave intensity at the shock, and the scalelength for wave decrease upstream (Kennel et al., 1983). It should be noted, however, that there exist more anisotropic events, normally in association with quasi-perpendicular shocks, whose signatures indicate that they may be energetic solar flare ions reaccelerated in a single encounter with the shock rather than ions accelerated out of the solar wind (Decker, 1981).

An interesting feature of the self-consistent theory is the inverse relationship between ion and wave intensity near the shock front on the one hand and the scalelength of their decrease with distance upstream on the other. That is, we have  $L(v) \sim A(v)/f(v,0)$  and  $I(k,0) \sim B(k) f(\Omega/k,0)$ , where  $L(v)$  is the scalelength for wave and ion intensity decrease away from the shock,  $f(v,0)$  is the ion omnidirectional distribution at the shock,  $I(k,0)$  is the wave intensity at the shock,  $\Omega$  is the proton gyrofrequency, and  $A(v)$  and  $B(k)$  are prescribed functions. This inverse relationship may explain why interplanetary traveling shocks at 1 AU have on average smaller energetic ion and hydromagnetic wave enhancements than Earth's bowshock but more extensive foreshocks. Characteristic "diffuse" ion and ULF wave (integrated over frequencies cyclotron resonant with the ions) energy densities just upstream of the bowshock are approximately  $350 \text{ eV-cm}^{-3}$  and  $40 \text{ eV-cm}^{-3}$ , and the characteristic scalelength for 30 keV protons is  $\sim 7 R_E$ . For a large interplanetary traveling shock event the corresponding numbers are approximately  $150 \text{ eV-cm}^{-3}$ ,  $30 \text{ eV-cm}^{-3}$  and  $50 R_E$ . The lower maximum upstream ion and wave intensities at interplanetary traveling shocks are probably due to smaller injection rates due to their smaller Mach numbers. Indeed the large variation in maximum ion intensities observed at interplanetary traveling shocks (Wenzel et al., 1983) is probably due to a sensitive dependence of ion injection rates on Mach numbers near the critical Mach number. Although, as we have discussed, it is possible to predict injection energies, unfortunately it is not yet possible to calculate reliably the actual injection rate.

A final noteworthy feature of shock acceleration in interplanetary space is the deceleration of the upstream plasma by the energetic ion pressure gradient. A decrease in solar wind speed of up to 20 km/s is observed upstream of Earth's bowshock (Bame et al., 1980; Bonifazi et al., 1980). The effect is small in interplanetary space where energetic ion pressures are limited by ion losses due to the spatial configuration of the shock or by limited acceleration time. However, the mediation of the shock by energetic ion pressure and the resulting reduction in compression at the subshock (which can limit ion injection) is of central importance in the shock acceleration of galactic cosmic rays by interstellar shocks where the cosmic ray pressure is comparable to all other pressures (Axford, 1981; McKenzie and Volk, 1982).

## VII. OUTSTANDING PROBLEMS

The following list summarizes the foremost outstanding problems, more or less in decreasing order of priority, that are expected to engage investigators of collisionless shock physics in the foreseeable future.

### Quasi-parallel Shocks

What is their full field and particle structure?

What role does leakage of downstream into upstream plasma play in their profile?

How are their intrinsic properties separated from those dependent on communication with other sections of the same shock?

### Quasi-perpendicular, turbulent shocks

How (by what instabilities) are ions actually thermalized initially in the shock ramp and the more completely behind it?

Is there an electrostatic subshock?

Is there a second, supercritical Mach number? If so, what is it, what does it depend on, and how does it effect shock structure?

Where are the sources of seed ions and electrons for the foreshock?

What role do multiple reflections of ions play in the reflection process?

### Quasi-perpendicular, laminar shocks

How are ions heated in the laminar shock?

Are standing waves stable? If not, what becomes of them?

### Parametric Structure

What is the full, definable, macroscopic, quantitative categorization of shock structures by , Mach number, etc.? What nomenclature should be standardized to describe them, e.g., quasi-laminar; quasi-turbulent?

### Ion Acceleration

How, if at all, do suprathermal ions at interplanetary shocks differ from those at the bow-shock?

### Electrons

What is the detailed origin of the flat-topped distributions?

What control does local geometry have on potential-drop heating of electrons?

### Waves

What is the origin of ion sound turbulence?

What role do nonlinear wave-wave interactions play in shock wave dissipation?

How, if at all, do nonlinear plasma waves participate in the thermalization process?

### Global Aspects

What is the large scale closure system of the shock currents?

Can (should?) structural factors be incorporated in the jump conditions in multifluid, MHD flow codes for planetary shocks?

To what extent does quasi-parallel/quasi-perpendicular geometry and its variability affect the overall solar wind-magnetosphere interaction?

## VIII. INVESTIGATE AVENUES

The study of collisionless shocks can, and has been, pursued by laboratory experimentation, theory/simulation, and natural observation. Recent years have seen an abundance of observational data, and a rising use of numerical simulations. In contrast, laboratory experimentation has been abandoned, and purely analytic theory has been substantially curtailed. It is clear that in the immediate future the predominance of observation and simulation will continue. Nevertheless, each of these approaches has its advantages and its limitations, and none should be excluded from future consideration. We summarize here the benefits of each.

### LABORATORY EXPERIMENTS

Laboratory devices suffer from wall effects, finite times of runs, poor approximations to collisionless densities and scaling defects in general, and difficulties in measurement of shock properties by noninvasive methods. Some of these deficiencies can be overcome simply by enlarging the apparatus, and the corresponding experimental budgets, which would still be small compared to those of typical spacecraft programs. The major advantage of laboratory operations, impossible to overemphasize, is control over, hence knowledge of, the input plasmas and the perturbations superposed on them. This means that experiments can be repeated at will and measurements can be made with almost arbitrary precision and reliability.

It is possible to put together quite a good picture of shock behavior by combining the results of piston and plasma-stream experiments. The former show the shocks in the early stage of evolution and permit precise measurements of structure and plasma conditions over a fairly wide parameter range. The plasma stream experiments display the characteristics of the fully-developed shock structure, but are restricted to a much narrower parameter range. The observation of the instability growing on the whistler may be an indication of the transition between the stationary and the turbulent states.

A piston-type experiment which permitted observation for several ion gyroperiods, thus combining the advantages of both types of experiments, would probably be the most fruitful way of obtaining more information on laboratory shocks. This would of necessity be quite large (several meters diameter) but could be designed so that magnetic scale lengths ( $c/\omega_{pe}$  and  $c/\omega_{pi}$ ) and electrostatic scale length ( $\lambda_D$ ) could be probed simultaneously.

## THEORY AND SIMULATION

Theoretical analysis is the symbolic shorthand whereby we communicate to ourselves and each other our understanding of plasma phenomena. In nonlinear regimes such as shocks, however, the symbolism can become as complicated as the plasma and can defy the kinds of traditional pencil and paper manipulation that contribute to understanding, or it can become so statistically general as to convey no understanding of details in real plasma. These difficulties can be overcome to some extent by conducting numerical "experiments" in which the motions of many particles and the values of fields governed by the fundamental equations are tracked in a large computer, starting with only relatively basic equations. Computer runs to stationarity in three-dimensional spaces, keeping track of phase relationships for two or more species of particles, can be slow, time consuming, voracious of computer resources, and therefore very expensive.

Nonetheless, simulation is an important link between theory and observations. Simulation can provide nonlocal, nonlinear, time-dependent solutions far beyond the usual capabilities of analytical theory (e.g., particle simulations as contrasted by Vlasov theory). Furthermore, information can also be extracted from the numerical calculations and put in a form which can be directly compared with observations (e.g., particle distribution functions). Thus, observation, theory and simulation are all strongly connected and mutually aid in the analysis and understanding of each other. As larger computers and more sophisticated numerical methods become available, we can look forward to more global simulations and thus to an even greater overall understanding of the bowshock and its relation to the rest of the magnetosphere.

Shock simulations are numerical solutions to the supersonic, super-Alfvénic plasma flow problem for given boundary conditions, geometries, and properties of the individual plasma species. The objective of a simulation is to model as closely as possible the observed properties of a shock, such as the Earth's bowshock, or the theoretically postulated properties, such as a relativistic shock from super novae. The complete simulation requires a 3-dimensional model which follows the evolution of the individual particles making up the various plasma species, and is large enough so that all plasma scale lengths, both in distance and in time, can be examined. Such a code is not economically feasible, so that simulations being employed today each represent partial solutions to the total shock problem. In what follows we will summarize briefly the methods currently being used, their strengths and weaknesses, and possibilities for their future use.

One dimensional hybrid codes assume that spatial variations occur only in the shock normal direction, that ions are treated kinetically (the individual ion orbits are followed) and that the electrons may be treated as a resistive or non-resistive fluid. The principle advantage of this method is that large systems (many ion inertial lengths) may be examined for long times (many ion gyroperiods). This method has been used to demonstrate the importance of reflected ions in

providing dissipation for perpendicular and quasi-perpendicular shocks (Leroy et al., 1981, 1982), and to generate ion velocity distributions which are similar to those observed near the Earth's bowshock (Paschmann et al., 1982). A disadvantage of the method is that electron dissipation must be provided through a phenomenological resistivity, as electron kinetic and cross-field driven plasma instabilities cannot be excited in a one-dimensional fluid model. Further, such non-Maxwellian features as an electron heat flux and temperature anisotropy cannot be examined using the fluid model. Future research using this method should include (i) adding a second spatial dimension and electron inertia to examine the effect of fluid cross-field driven instabilities, (ii) a detailed study of the acceleration and scattering of ions both at an oblique shock and in the region of upstream turbulence (Leroy and Winske, 1983; Kan and Swift, 1983), and (iii) the effect of electron resistivity (Leroy and Winske, 1983), and long wavelength turbulent ion waves on oblique shock structure (Kan and Swift, 1983).

One and two dimensional particle simulation codes (Brackbill and Forslund, 1982) treat both electron and ions kinetically, thus retaining all particle information of both species; however, the size of the simulation system and the length in time of the run must be made small, a consequence of the increased computational cost. The primary advantage of this method is that it can be used to identify the self-consistent dissipation mechanisms present in collisionless shocks. For example, the method has shown that the lower-hybrid-drift instability is a source of dissipation in quasi-perpendicular shocks (Forslund et al., 1983), and that turbulent ion heating by the back-scatter whistler decay instability (Forslund et al., 1983), could be important in quasi-parallel shocks (Quest et al., 1983). Further, simulations show a resistive electron response for quasi-perpendicular shocks and an adiabatic response for quasi-parallel shocks, thus lending support to the resistive hybrid simulation method in the former case and casting some doubts in the latter. Topics for the future using the particle method include (i) examining in detail the sources of dissipation at the shock as a function of shock normal angle and plasma parameters, (ii) determining in what parameter space the one-dimensional and hybrid codes agree with the two-dimensional particle codes, and (iii) examining the evolution of the electron species.

In summary, hybrid simulations are useful for describing the long-scale length, low frequency evolution of the shock, while particle simulations are used to determine the self-consistent mechanisms of shock dissipation and to provide a check on the results of the hybrid codes.

## SPACECRAFT OBSERVATIONS

The benefit of direct observations by spacecraft where it matters, namely in natural shocks themselves, needs no elaboration, but it is important to note the drawbacks of relying on satellite measurements. There is virtually no control, beyond selective analysis of data, of the conditions or states of the plasmas or shocks under investigation. In addition, the measurements must be carried out, at great expense, in remote locations after full control, let alone inspection, of the measuring apparatus and its data handling machinery has been relinquished. Finally, comprehensive observations can only be carried out by numerous investigators, each responsible for a part of the picture, with the results dependent on cooperation among all concerned, many of whom are distant from each other. Despite these problems, satellite experiments offer a view into a surprising range of plasma shock conditions that occur in the solar terrestrial environment.

Figure 10-16 compares scalelengths characteristic of a nominal interplanetary plasma (Feldman et al., 1977) with those actually observed in terrestrial bowshocks, and also shows the relationship of spacecraft characteristics to these scalelengths. Satellite dimensions are typically small enough that the spatial resolution of spacecraft observations should be adequate to resolve any of the pertinent scalelengths.

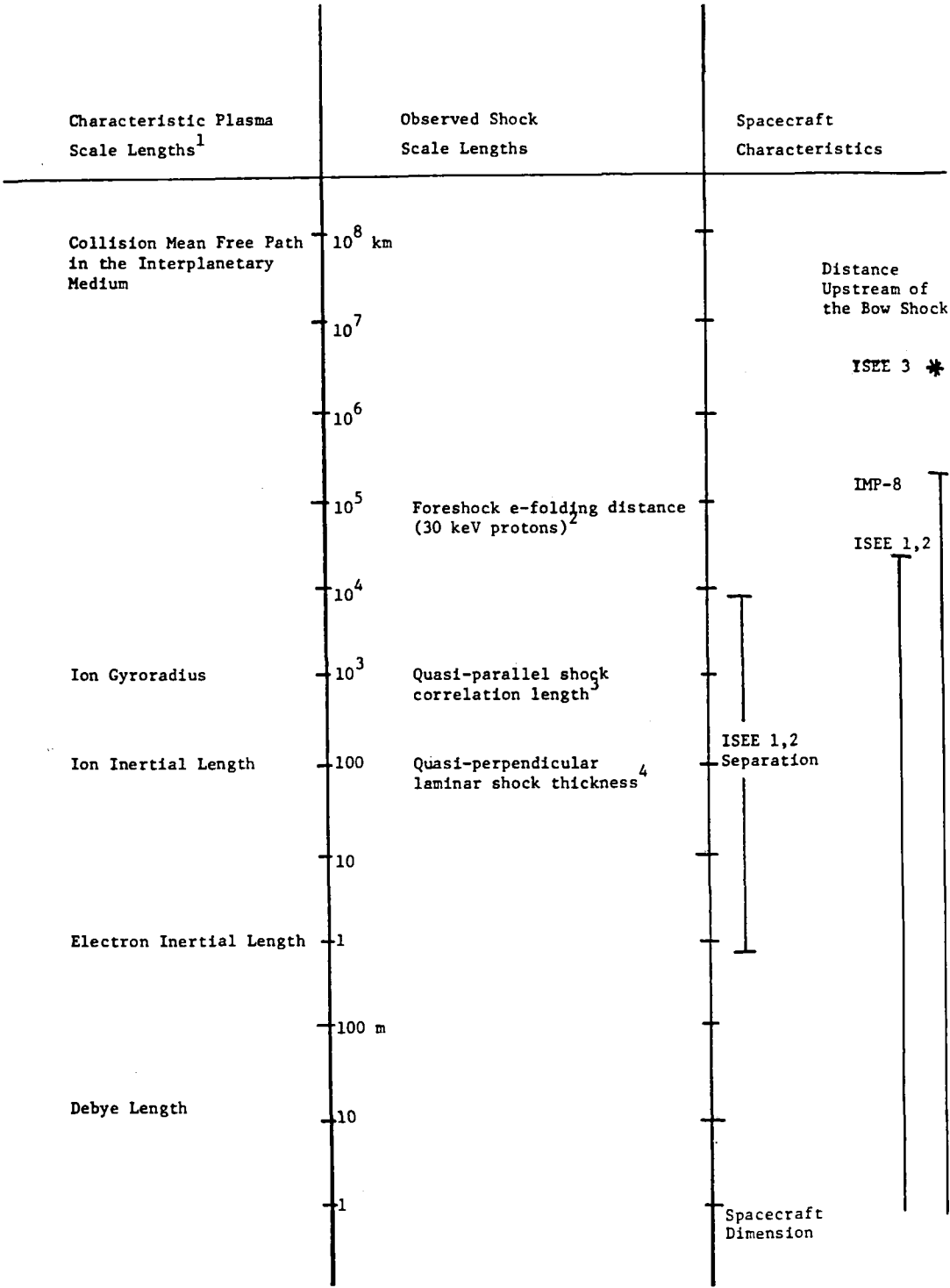


Figure 10-16. Relative scales of important plasma quantities, shock dimensions, and spacecraft parameters.

There is, however, in practice the difficulty that shock velocities with respect to the spacecraft are often too high to permit detailed measurements within the shock per se. This is illustrated in Table 10-1 where sampling rates of a range of currently operating instruments are listed (examples taken from IEEE Geosciences issue on ISEE instrumentation). The table includes a computation of the number of data points which would be obtained by each instrument in passing through a 100 km thick shock whose velocity was 17 km/sec (typical values from Russell et al.'s 1982 study of quasi-perpendicular laminar shocks). Field and plasma wave data rates are clearly adequate to resolve the details of the structure of such shocks. On the other hand, although the evaluation of plasma distribution functions within a given shock range has been studied in a number of special cases, current sampling rates are marginal for detailed study of most perpendicular and/or quasi-perpendicular shocks. Shock motion also contributes to the difficulty of distinguishing spatial and temporal effects, an ambiguity which has only recently been effectively addressed with the launch of the ISEE spacecraft. The range of ISEE-1, -2 separations, which brackets the scale of solar wind ion motions, has proven well suited to the study of both bowshocks and interplanetary shock waves. In addition research using ISEE-1 and -2 in combination with IMP-8 and ISEE-3 has proven very useful in attempts to map the large scale foreshock regions in which wave and particle effects extend upstream of shocks themselves.

**TABLE 10-1**  
**CURRENT (ISEE) SAMPLING RATES**  
**FOR SHOCK RELATED PLASMA PARAMETERS**

<u>QUANTITY</u>	<u>HIGHEST SAMPLING RATE</u>	<u>NO. SAMPLES/ 6 SEC SHOCK</u>	<u>INSTRUMENT</u>
Electric Field (2D)	32/Sec	192	DC Electric Field
Magnetic Field (3D)	16/Sec	96	Fluxgate Magnetometer
<b>Electrons</b>			
Density	32/Sec	192	Propagation Experiment
	16/Sec	96	Sounder
$F_E$ (V)	3/Sec	2	Electron Spectrometer
<b>Protons</b>			
Solar Wind	12/Sec	-	Solar Wind Instrument
$F_P$ (V) (2D)	3/Sec	2	Fast Plasma Experiment
Plasma Waves	4 Spectra/Sec	24	Plasma Wave



In Figure 10-17 typical bowshock parameters are presented within the context of the shock classification scheme proposed by Greenstadt (1974) and Formisano (1977). This scheme, which applied to quasi-perpendicular shocks, divides shocks into four classes depending on  $\beta$  and Mach number. Numbers following the designations represent estimates of the percentage of actual bowshocks which fall in each category. The small box labelled ISEE encloses mean solar wind conditions at ISEE + one standard deviation and the outer dotted line encloses values within the 5-95% confidence level (solar wind values from Feldman et al., 1977). The box labelled 'Interplanetary shocks' represents values associated with the set of shocks studied by Russell et al. (1983).

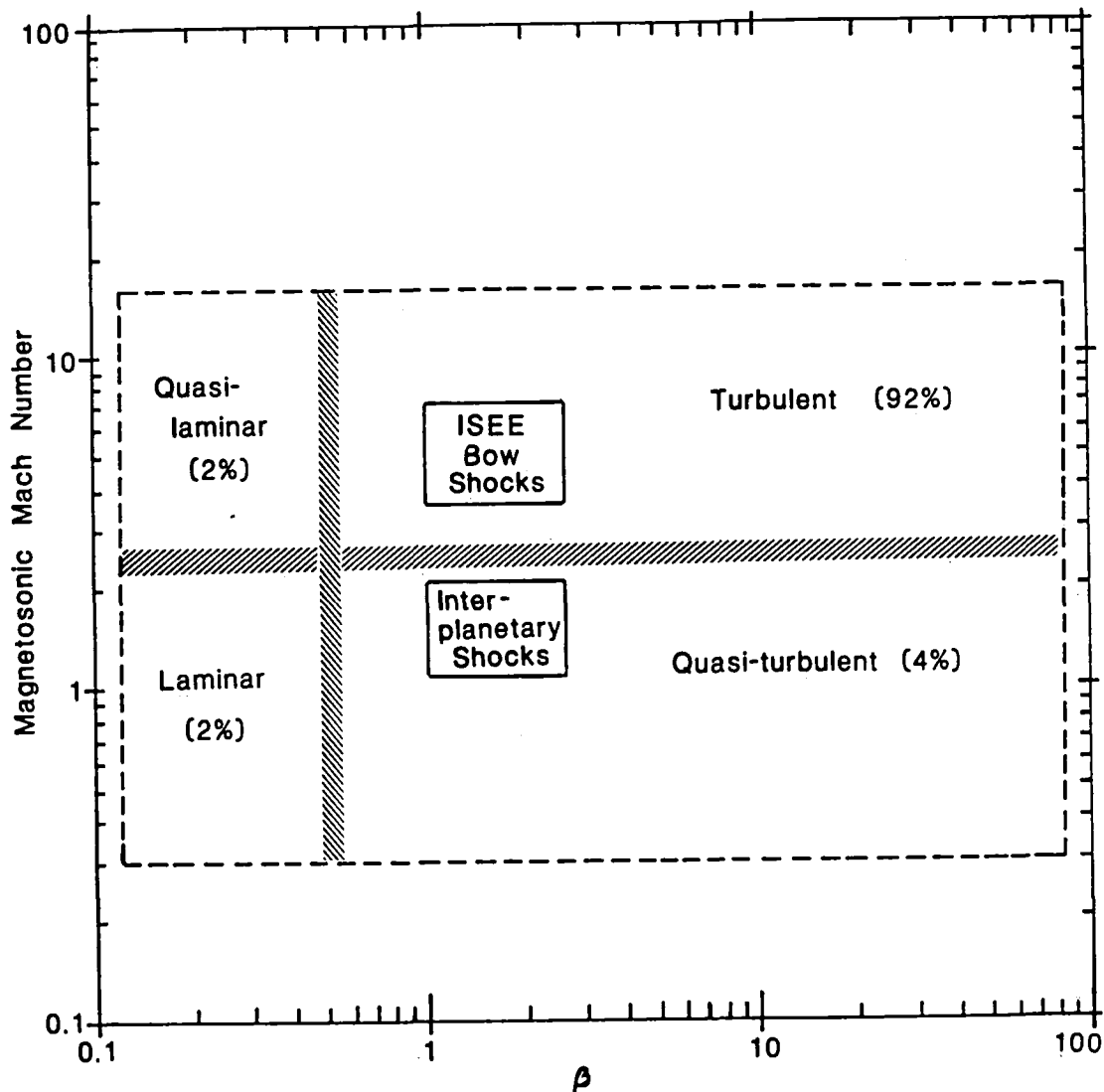


Figure 10-17. Relative percentage of occurrences of subcategories of quasi-perpendicular shock conditions dependent on Mach number and beta.

## IX. REFERENCES

- Anderson, K. A., R. P. Lin, F. Martel, C. S. Lin, G. K. Parks, and H. Reme, Thin sheets of energetic electrons upstream from the Earth's bow shock, Geophys. Res. Lett., **6**, 401, 1979.
- Anderson, R. R., G. K. Parks, T. E. Eastman, D. A. Gurnett, and L. A. Frank, Plasma waves associated with energetic particles streaming into the solar wind from the Earth's bow shock, J. Geophys. Res., **86**, No. A6, 4493-4510, 1981.
- Asbridge, J. R., S. J. Bame, and I. B. Strong, Outward flow of protons from the earth's bow shock, J. Geophys. Res., **73**, 5777, 1968.
- Auer, P. L., and W. H. Evers, Jr., Collision-free shock formation in finite temperature plasmas, Phys. Fluids, **14**, 1177, 1971.
- Auer, R. -D., and H. J. Volk, Parallel, high B shocks and relaxation phenomena, Astrophys. & Space Sci., **22**, 243, 1973.
- Auer, P. L., H. Hurwitz, Jr., and R. W. Kilb, Large amplitude magnetic compression of a collision-free plasma, 2, Development of a thermalized plasma, Phys. Fluids, **5**, 298, 1962.
- Auer, P. L., R. W. Kilb, and W. F. Crevier, Thermalization of the earth's bow shock, J. Geophys. Res., **76**, 2927, 1971.
- Axford, W. I., Acceleration of cosmic rays by shock waves, Proc. Int. Conf. Cosmic Rays 17th, **12**, 155, 1981.
- Bame, S. J., J. R. Asbridge, J. T. Gosling, M. Halbig, G. Paschmann, N. Sckopke, and H. Rosenbauer, High temporal resolution observations of electron heating at the bow shock, Space Sci. Rev., **23**, 75, 1979.
- Bame, S. J., J. R. Asbridge, W. C. Feldman, J. T. Gosling, G. Paschmann, and N. Sckopke, Deceleration of the solar wind upstream from the earth's bow shock and the origin of diffuse upstream ions, J. Geophys. Res., **85**, 2981, 1980.
- Barnes, A., Theory of generation of bow-shock-associated hydromagnetic waves in the upstream interplanetary medium, Cosmic Electrodyn., **1**, 90, 1970.
- Biskamp, D., Collisionless shock waves in plasmas, Nuclear Fusion, **13**, 719, 1973.
- Biskamp, D., and H. Welter, Numerical studies of magnetosonic collisionless shock waves, Nucl. Fusion, **12**, 633, 1972a.
- Biskamp, D., and H. Welter, Ion heating in high-mach number, oblique collisionless shock waves, Phys. Rev. Lett., **28**, 410, 1972b.
- Bonifazi, C. and G. Moreno, Reflected and diffuse ions backstreaming from the earth's bow shock, 1, basic properties, J. Geophys. Res., **86**, 4397, 1981a.
- Bonifazi, C., and G. Moreno, Reflected and diffuse ions backstreaming from the earth's bow shock, 2, origin, J. Geophys. Res., **86**, 4405, 1981b.

- Bonifazi, C., G. Moreno, A. J. Lazarus, and J. D. Sullivan, Deceleration of the solar wind in the earth's foreshock region: ISEE 2 and IMP 8 observations, J. Geophys. Res., **85**, 6031, 1980.
- Brackbill, J. U., and D. W. Forslund, An implicit method for electromagnetic plasma simulation in two dimensions, J. Comp. Res., **46**, 271, 1982.
- Burlaga, L. F., and K. W. Ogilvie, Observations of the magnetosheath-solar wind boundary, J. Geophys. Res., **73**, 6167, 1968.
- Chao, J. K., and M. J. Wiskerchen, The ratio of specific heats for post-shock plasmas of a detached bow shock: An MHD model, J. Geophys. Res., **79**, 4769, 1974.
- Chodura, R., A hybrid fluid particle model of ion heating in high mach-number shock waves, Mech. Fusion, **15**, 55, 1975.
- Cipolla, J. W., K. I. Golden, and M. B. Silevich, Ion cyclotron beam mode-whistler mode plasma instabilities and their role in parallel shock structures, Phys. Fluids, **20**, 282, 1977.
- Craig, A. D., S. Nakai, D. D. R. Summers, and J. W. M. Paul, Measurement of long wavelength turbulence within a collisionless shock by scattering of radiation from a CO<sub>2</sub> laser, Phys. Rev. Lett., **32**, 975, 1974.
- Crooker, N. U., T. E. Eastman, and G. S. Stiles, Observations of plasma depletion in the magnetosheath at the dayside magnetopause, J. Geophys. Res., **84**, 869, 1979.
- Decker, R. B., The modulation of low-energy proton distributions by propagating interplanetary shock waves: a numerical simulation, J. Geophys. Res., **86**, 4537, 1981.
- Decker, G. and A. E. Robson, Instability of the whistler structure of oblique hydromagnetic shocks, Phys. Rev. Lett., **29**, 1071, 1972.
- Diodato, L., E. W. Greenstadt, G. Moreno, and V. Formisano, A statistical study of the upstream wave boundary outside the Earth's bow shock, J. Geophys. Res., **81**, 199, 1976.
- Dobrowolny, M., and V. Formisano, The structure of the earth's bow shock, Rev. Nuovo Cim, **3**, 419, 1973.
- Dryer, M., and G. R. Heckman, On the hypersonic analogue as applied to planetary interactions with the solar plasma, Planet. Space Sci., **15**, 515, 1967.
- Eastman, T. E., R. R. Anderson, L. A. Frank, and G. K. Parks, Upstream particles observed in the Earth's foreshock region, J. Geophys. Res., **86**, 4379, 1981.
- Edmiston, J. P., C. F. Kennel, and D. Eichler, Escape of heated ions upstream of quasi-parallel shocks, Geophys. Res. Lett., **9**, 531, 1982.
- Eselevich, V. G., Shock-wave structure in collisionless plasmas from results of laboratory experiments, Space Sci. Rev., **32**, 65, 1982.

- Fairfield, D. H., Bow shock associated waves observed in the far upstream interplanetary medium, J. Geophys. Res., **74**, 3541, 1969.
- Fairfield, D. H., Average and unusual positions of the earth's magnetopause and bow shock, J. Geophys. Res., **76**, 6700, 1971.
- Fairfield, D. H., Magnetic fields of the magnetosheath, Rev. Geophys. Space Sci., **14**, 117, 1976.
- Fairfield, D. H., Whistler waves observed upstream from collisionless shock, J. Geophys. Res., **79**, 1368, 1974.
- Feldman, W. C., J. R. Asbridge, S. J. Bame, and J. T. Gosling, Plasma and magnetic fields from the sun, O. K. White, editor, The Solar Output and its Variation, Colo. Assoc. Univ. Press, Boulder, Colo., 1977.
- Feldman, W. C., S. J. Bame, S. P. Gary, J. T. Gosling, D. J. McComas, M. F. Thomsen, G. Paschmann, N. Sckopke, M. M. Hoppe, and C. T. Russell, Electron heating within the earth's bow shock, Phys. Rev. Lett., **49**, 199, 1982.
- Feldman, W. C., R. C. Anderson, S. J. Bame, S. P. Gary, J. T. Gosling, D. J. McComas, M. F. Thomsen, G. Paschmann, and M. M. Hoppe, Electron velocity distributions near the earth's bow shock, J. Geophys. Res., **88**, 96, 1983a.
- Feldman, W. C., R. C. Anderson, S. J. Bame, J. T. Gosling, R. D. Zwickl, and E. J. Smith, Electron velocity distributions near the interplanetary shocks, J. Geophys. Res., in preparation, 1983b.
- Formisano, V., The physics of the earth's collisionless shock wave, J. Physique, **38**, C6-55-C6-88, 1977.
- Formisano, V., Orientation and shape of the earth's bow shock in three dimensions, Planet. Space Sci., **27**, 1151, 1979a.
- Formisano, V., Low-frequency waves observed in the vicinity of the earth's bow shock, Nuovo Cimento, **20**, 789, 1979b.
- Formisano, V., Plasma processes at collisionless shock waves, Proc. Int. Sch. Wkshp in Plasma Astrophys., Vienna, ESA SP-161, 1981.
- Formisano, V., and P. Hedgecock, Solar wind interaction with the earth's magnetic field, 3, On the earth's bow shock structure, J. Geophys. Res., **78**, 3745, 1973a.
- Formisano, V., and P. Hedgecock, On the structure of the turbulent bow shock, J. Geophys. Res., **78**, 6522, 1973b.
- Formisano, V., and R. Torbert, Ion acoustic wave forms generated by ion-ion streams at the earth's bow shock, Geophys. Res. Lett., **9**, 207, 1982.
- Formisano, V., C. T. Russell, J. D. Means, E. W. Greenstadt, F. L. Scarf, and M. Neugebauer, Collisionless shock waves in space: a very high-beta structure, J. Geophys. Res., **80**, 2030, 1975.

- Forslund, D. W., and J. P. Friedberg, Theory of laminar collisionless shocks, Phys. Rev. Lett., **27**, 1189, 1971.
- Forslund, D. W., J. M. Kindel, and E. L. Lindman, Parametric excitation of electromagnetic waves, Phys. Rev. Lett., **29**, 249, 1972.
- Forslund, D. W., K. B. Quest, J. U. Brackbill, and K. Lee, Collisionless dissipation in quasi-perpendicular shocks, J. Geophys. Res., **88**, in press, 1983.
- Fredricks, R. W., F. V. Coroniti, C. F. Kennel, and F. L. Scarf, Fast time-resolved spectra of electrostatic turbulence in the earth's bow shock, Phys. Rev. Lett., **24**, 994, 1970.
- Galeev, A. A., Collisionless shocks, Physics of Solar Planetary Environment, edited by D. J. Williams, American Geophysical Union 464, 1976.
- Gary, S. P., J. T. Gosling, and D. W. Forslund, The electromagnetic ion beam instability upstream of the earth's bow shock, J. Geophys. Res., **86**, 6691, 1981.
- Gloeckler, G., F. M. Ipavich, C. Y. Fan, and D. Hovestadt, Post-shock spikes: a new feature of proton and alpha enhancements associated with an interplanetary shock wave, Geophys. Res. Lett., **1**, 65, 1974.
- Gosling, J. T., J. R. Asbridge, S. J. Bame, and I. B. Strong, Vela 2 measurements of magnetopause and bow shock positions, J. Geophys. Res., **72**, 101, 1967.
- Gosling, J. T., J. R. Asbridge, S. J. Bame, G. Paschmann, and N. Sckopke, Observations of two distinct populations of bow shock ions, Geophys. Res. Lett., **5**, 957, 1978.
- Gosling, J. T., J. R. Asbridge, S. J. Bame, W. C. Feldman, R. D. Zwickl, G. Paschmann, N. Sckopke, and R. J. Hynds, Interplanetary ions during an energetic storm particle event: the distribution function from solar wind thermal energies to 1.6 MeV, J. Geophys. Res., **86**, 547, 1981.
- Gosling, J. T., M. F. Thomsen, S. J. Bame, W. C. Feldman, G. Paschmann, and N. Sckopke, Evidence for specularly reflected ions upstream from the quasi-parallel bow shock, Geophys. Res. Lett., **9**, 1333, 1982.
- Gosling, J. T., S. J. Bame, W. C. Feldman, G. Paschmann, N. Sckopke, and C. T. Russell, Suprathermal ions upstream from interplanetary shocks, J. Geophys. Res., Submitted, 1983.
- Greenstadt, E. W., Binary index for assessing local bow shock obliquity, J. Geophys. Res., **77**, 5467, 1972.
- Greenstadt, E. W., Structure of the terrestrial bow shock, Solar Wind Three, edited by C. T. Russell, Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA, 1974.
- Greenstadt, E. W., and R. W. Fredricks, Plasma instability modes related to the Earth's bow shock, magnetospheric physics, edited by B. M. McCormac, D. Reidel, Dordrecht, Holland, 355, 1974.

- Greenstadt, E. W., and R. W. Fredricks, Shock systems in collisionless space plasmas, in solar system plasma physics, vol. III, edited by L. J. Lanzerotti, C. F. Kennel, and E. N. Parker, p. 3, North-Holland, Amsterdam, 1979.
- Greenstadt, E. W., I. M. Green, G. T. Inouye, A. J. Hundhausen, S. J. Bame, and I. B. Strong, Correlated magnetic field and plasma observations of the earth's bow shock, J. Geophys. Res., **73**, 51, 1968.
- Greenstadt, E. W., I. M. Green, G. T. Inouye, D. S. Colburn, J. H. Binsack, and E. F. Lyon, Dual satellite observation of the Earth's bow shock, 1, the thick pulsation shock, 2, field-aligned upstream waves, 3, field-determined shock structure, cosmic electrodyn., **1**, 279, 1970.
- Greenstadt, E. W., C. T. Russell, F. L. Scarf, V. Formisano, P. Hedgecock, M. Neugebauer, and R. E. Holzer, Structure of quasi-parallel, quasi-laminar bow shock, J. Geophys. Res., **82**, 651, 1977.
- Greenstadt, E. W., C. T. Russell, and M. Hoppe, Magnetic field orientation and suprathermal ion streams in the earth's foreshock, J. Geophys. Res., **85**, 3473, 1980a.
- Greenstadt, E. W., C. T. Russell, J. T. Gosling, S. J. Bame, G. Paschmann, G. K. Parks, K. A. Anderson, F. L. Scarf, R. R. Anderson, D. A. Gurnett, R. P. Lin, C. S. Lin, and H. Reme, A macroscopic profile of the typical quasi-perpendicular bow shock: ISEE 1 and -2, J. Geophys. Res., **85**, 2124, 1980b.
- Greenstadt, E. W., R. W. Fredricks, C. T. Russell, F. L. Scarf, R. R. Anderson, and D. A. Gurnett, Whistler mode wave propagation in the solar wind near the bow shock, J. Geophys. Res., **86**, 4511-4516, 1981.
- Greenstadt, E. W., M. M. Hoppe, and C. T. Russell, Large-amplitude magnetic variations in quasi-parallel shocks: Correlation lengths measured by ISEE-1 and -2, Geophys. Res. Lett., **9**, 781-784, 1982.
- Guha, J. K., D. L. Judge, and J. H. Marburger, OGO 5 magnetic-field data near the Earth's bow shock: a correlation with theory, J. Geophys. Res., **77**, 604, 1972.
- Gurgiolo, C., G. K. Parks, B. H. Mauk, C. S. Lin, K. A. Anderson, R. P. Lin, and H. Reme, Non ExB ordered ion beams upstream of the earth's bow shock, J. Geophys. Res., **86**, 4415, 1981.
- Gurnett, D. A., R. R. Anderson, F. L. Scarf, R. W. Fredricks, and E. J. Smith, Initial results from the ISEE-1 and -2 plasma wave investigation, Space Sci. Rev., **23**, 103, 1979.
- Heppner, J. P., M. Sugiura, T. L. Skillman, B. G. Ledley, and M. Campbell, OGO-A magnetic-field observations, J. Geophys. Res., **72**, 5417, 1967.
- Hoppe, M. M. and C. T. Russell, Particle acceleration at planetary bow shock waves, Nature, **295**, 41, 1982.
- Hoppe, M. M., and C. T. Russell, Plasma rest frame frequencies and polarizations of the low frequency upstream waves: ISEE 1 and 2 observations, J. Geophys. Res., **88**, 2021, 1982.

- Hoppe, M., C. T. Russell, L. A. Frank, T. E. Eastman, and E. W. Greenstadt, Upstream hydromagnetic waves and their association with backstreaming ion populations: ISEE-1 and -2 observations, J. Geophys. Res., **86**, 4471, 1981.
- Howe, H. C., Jr., Explorer 33 and 35 plasma observations of the interaction region between the solar wind and the magnetic field of the earth, Ph. D. Dissertation, MIT, 1971.
- IEEE Trans. Geosci. Electronics, GE-16, 151-270, 1978.
- Ipavich, F. M., A. B. Galvin, G. Gloeckler, M. Scholer, and D. Hovestadt, A statistical survey of ions observed upstream of the earth's bow shock: energy spectra, composition, and spatial variation, J. Geophys. Res., **86**, 4337, 1981a.
- Ipavich, F. M., M. Scholer, and G. Gloeckler, Temporal development of composition, spectra, and anisotropies during upstream particle events, J. Geophys. Res., **86**, 11153, 1981b.
- Kan, J. R., and D. W. Swift, Structures of quasi-parallel bow shock: results of numerical simulations, J. Geophys. Res., **88**, 6919, 1983.
- Keilhacker, M., and K. H. Steuer, Time-resolved light-scattering measurements of the spectrum of turbulence within a high- $\beta$  collisionless shock wave, Phys. Rev. Lett., **26**, 694, 1971.
- Kellogg, P. J., Shock waves in interplanetary plasma, Univ. Minn. Tech. Rep. CR-69, Feb. 1964. (Also Proceedings of International Conference on Cosmic Rays, Jaipur, India, 1963.)
- Kennel, C. F., Collisionless shocks and upstream waves and particles: introductory remarks, J. Geophys. Res., **86**, 4325, 1981.
- Kennel, C. F., and H. E. Petschek, Limit on stably trapped particle fluxes, J. Geophys. Res., **71**, 1, 1966.
- Kennel, C. F., and R. Z. Sagdeev, Collisionless shock waves in high B plasmas, 1, J. Geophys. Res., **72**, 3303, 1967.
- Kennel, C. F., F. L. Scarf, F. V. Coroniti, E. J. Smith, and D. A. Gurnett, Non-local plasma turbulence associated with interplanetary shocks, J. Geophys. Res., **87**, 1774, 1982.
- Kennel, C. F., F. L. Scarf, F. V. Coroniti, C. T. Russell, E. J. Smith, B. T. Tsurutani, K. P. Wenzel, R. Reinhard, T. R. Sanderson, W. C. Feldman, G. K. Parks, F. S. Mozer, M. A. Temerin, and R. R. Anderson, Plasma and energetic particle structure of a collisionless quasi-parallel shock, J. Geophys. Res., submitted for publication, 1983.
- Leboeuf, J. N., T. Tajima, C. F. Kennel, and J. M. Dawson, Global simulation of the time dependent magnetosphere, Geophys. Res. Lett., **5**, 609, 1978.
- Lee, M. A., Coupled hydromagnetic waves excitation and ion acceleration upstream of the earth's bow shock, J. Geophys. Res., **87**, 5063, 1982.
- Lee, M. A., Coupled hydromagnetic wave excitation and ion acceleration at interplanetary travelling shocks, J. Geophys. Res., in press, 1983.

- Lemons, D. S., and S. P. Gary, Current-driven instabilities in a laminar perpendicular shock, J. Geophys. Res., **83**, 1625, 1978.
- Lemons, D. S., J. R. Asbridge, S. J. Bame, W. C. Feldman, S. P. Gary, and J. T. Gosling, The source of electrostatic fluctuations in the solar wind, J. Geophys. Res., **84**, 2135, May 1, 1979.
- Leroy, M. M., and D. Winske, Backstreaming ions from oblique earth bow shocks, Annales Geophysicae, **1**, in press, 1983.
- Leroy, M. M., C. C. Goodrich, D. Winske, C. S. Wu, and K. Papadopoulos, Simulations of a perpendicular bow shock, Geophys. Res. Lett., **8**, 1269, 1981.
- Leroy, M. M., D. Winske, C. C. Goodrich, C. S. Wu, and K. Papadopoulos, The structure of perpendicular bow shocks, J. Geophys. Res., **87**, 5081, 1982.
- Lyon, J., S. H. Brecht, J. A. Fedder, and P. Palmadesso, The effects on the earth's magnetotail from shocks in the solar wind, Geophys. Res. Lett., **7**, 721, 1980.
- Mason, R. J., Ion and electron pressure effects on magnetosonic shock formation, Phys. Fluids, **15**, 1082, 1972.
- McKenzie, J. F., and H. J. Volk, Non-linear theory of cosmic ray shocks including self-generated Alfvén waves, Astron. Astrophys., **116**, 191, 1982.
- Mellott, M. M., and E. W. Greenstadt, ISEE-1 and -2 observations of laminar bow shocks: whistler precursors and shock structure, submitted to J. Geophys. Res., 1984.
- Michel, F. C., Detectability of disturbances in the solar wind, J. Geophys. Res., **70**, 1, 1965.
- Mihalov, J. D., J. H. Wolfe, and D. S. Intrilligator, Pioneer Venus observations of the solar wind - Venus interactions, J. Geophys. Res., **85**, 7613, 1980.
- Montgomery, M. D., J. R. Asbridge, and S. J. Bame, Vela 4 plasma observations near the earth's bow shock, J. Geophys. Res., **75**, 1217, 1970.
- Morse, D. L., and W. W. Destler, Laboratory study of high- $\beta$  plasma shock waves, Plasma Physics, **14**, 153, 1972.
- Ogilvie, K. W., J. D. Scudder, and M. Sugiura, Electron energy flux in the solar wind, J. Geophys. Res., **76**, 8165, 1971.
- Olbert, S. and B. Rossi, Introduction to the Physics of Space, McGraw Hill, New York, 1970.
- Papadopoulos, K., C. E. Wagner, and I. Haber, High-Mach-number turbulent magnetosonic shocks, Phys. Rev. Lett., **27**, 982, 1971.
- Paschmann, G., N. Sckopke, S. J. Bame, J. R. Asbridge, J. T. Gosling, C. T. Russell, and E. W. Greenstadt, Association of low-frequency waves with suprathermal ions in the upstream solar wind, Geophys. Res. Lett., **6**, 209, 1979.
- Paschmann, G., N. Sckopke, J. R. Asbridge, S. J. Bame, and J. T. Gosling, Energization of solar wind ions by reflection from the earth's bow shock, J. Geophys. Res., **85**, 4689, 1980.



- Paschmann, G., N. Sckopke, I. Papamastorakis, J. R. Asbridge, S. J. Bame, and J. T. Gosling, Characteristics of reflected and diffuse ions upstream from the earth's bow shock, J. Geophys. Res., **86**, 4355, 1981.
- Paschmann, G., N. Sckopke, S. J. Bame, and J. T. Gosling, Observations of gyrating ions in the foot region of a nearly perpendicular bow shock, Geophys. Res. Lett., **9**, 881, 1982.
- Paul, J. W. M. et al., Plasma physics and controlled thermonuclear research 1971 (Conference Proceedings, Madison, 17-23 June 1971) **3**, 251, (IAEA Vienna) 1971.
- Phillips, P. E. and A. E. Robson, Influence of reflected ions on the magnetic structure of a collisionless shock front, Phys. Rev. Lett., **29**, 154, 1972.
- Priest, E. R., and J. J. Sanderson, Ion acoustic instability in collisionless shocks, Plasma Phys., **14**, 951, 1972.
- Quest, K. B., D. W. Forslund, J. U. Brackbill, and K. Lee, Collisionless dissipation processes in quasi-parallel shocks, Geophys. Res. Lett., in press, 1983.
- Robson, A. E., Experiments on oblique shock waves, in collisionless shocks in the laboratory and space (Proc. ESRIN Study Group, Frascati, June 11-20, 1969) ESRO SP-51, p. 159, 1969.
- Robson, A. E., and J. Sheffield, Plasma physics and controlled thermonuclear research (Conference Proceedings, Novosibirsk, 1-7 August, 1968), **1**, 119, (IAEA Vienna) 1969.
- Rodriguez, P., and D. A. Gurnett, Electrostatic and electromagnetic turbulence associated with the earth's bow shock, J. Geophys. Res., **80**, 19, 1975.
- Rodriguez, P., and D. A. Gurnett, Correlation of bow shock plasma wave turbulence with solar wind parameters, J. Geophys. Res., **81**, 1871, 1976.
- Russell, C. T., and E. W. Greenstadt, Initial ISEE magnetometer results: shock observations, Space Sci. Rev., **23**, 3, 1979.
- Russell, C. T., and E. W. Greenstadt, Plasma boundaries and shocks, Rev. Geophys. Space Phys., **21**, 449, 1983.
- Russell, C. T., M. M. Hoppe, and W. A. Livesey, Overshoots in planetary bow shocks, Inst. Publ. 2182, Inst. of Geophys. and Planet. Phys., Univ. of Calif., Los Angeles, 1981.
- Russell, C. T., M. M. Hoppe, W. A. Livesey, J. T. Gosling, and S. J. Bame, ISEE-1 and -2, observations of laminar bow shocks: velocity and thickness, Geophys. Res. Lett., **9**, 1171, 1982.
- Russell, C. T., E. J. Smith, B. J. Tsurutani, J. T. Gosling, and S. J. Bame, Multiple spacecraft observations of interplanetary shocks: characteristics of the upstream ULF turbulence, Sol. Wind Proc. Conf., **5**, in press, 1983.
- Sagdeev, R. Z., Cooperative phenomena and shock waves, collisionless plasma, in Reviews of Plasma Physics, **4**, edited by M. A. Leontovich, Consultants Bureau, N.Y., 1966.

- Schwartz, S. J., M. F. Thomsen, and J. T. Gosling, Ions upstream of the earth's bow shock: a theoretical comparison of alternative source populations, J. Geophys. Res., **88**, 2039, 1983.
- Scholer, M., F. M. Ipavich, G. Gloeckler, D. Hovestadt, and B. Klecker, Pitch angle distributions of energetic protons near the earth's bow shock, Geophys. Res. Lett., **6**, 707, 1979.
- Scholer, M., F. M. Ipavich, G. Gloeckler, and D. Hovestadt, Acceleration of low-energy protons and alpha particles at interplanetary shock waves, J. Geophys. Res., **88**, 1977, 1983.
- Sckopke, N., G. Paschmann, S. J. Bame, J. T. Gosling, and C. T. Russell, Evolution of ion distributions across the nearly perpendicular bow shock: specularly and non-specularly reflected ions, J. Geophys. Res., **88**, 6121-6136, 1983.
- Scudder, J. D., and C. C. Goodrich, The energy gain of plasma electrons at collisionless shock waves, J. Geophys. Res., submitted, 1983.
- Scudder, J. D., D. L. Lind, and K. W. Ogilvie, Electron observations in the solar wind and magnetosheath, J. Geophys. Res., **78**, 6535, 1973.
- Scudder, J. D., A. Mangeney, C. S. Wu, C. Lacombe, G. Paschmann, J. T. Gosling, C. T. Russell, R. A. Anderson, P. Rodriguez, C. Harvey, and K. W. Ogilvie, Resolution of a high  $\beta$ , super critical quasi-perpendicular magnetosonic shock layer, J. Geophys. Res., to be submitted, 1984.
- Sentman, D. D., J. P. Edmiston, and L. A. Frank, Instabilities of low-frequency, parallel propagating electromagnetic waves in the earth's foreshock region, J. Geophys. Res., **86**, 7487, 1981.
- Sentman, D. D., M. F. Thomsen, S. P. Gary, W. C. Feldman, and M. M. Hoppe, The oblique whistler instability in the earth's foreshock, J. Geophys. Res., **88**, 2048, 1983.
- Shea, M. A., D. F. Smart, S. T. Wu, and S. Pinter, editors, Shock waves in solar corona and interplanetary space, Space Sci. Rev., **32**, pp. 1-272, 1982.
- Slavin, J. A., Average location of the terrestrial bow shock: 1963 - 1983, in preparation, 1983.
- Slavin, J. A., and R. E. Holzer, Solar wind flow about the terrestrial planets, 1, Modeling bow shock position and shape, J. Geophys. Res., **86**, 11,401, 1981.
- Slavin, J. A., R. S. Elphic, C. T. Russell, F. L. Scarf, J. H. Wolfe, J. D. Mihalov, D. S. Intrilligator, L. H. Brace, H. A. Taylor, Jr., and R. E. Daniel, Jr., The solar wind interaction with Venus: pioneer venus observations of bow shock location and structure, J. Geophys. Res., **85**, 7625, 1980.
- Slavin, J. A., R. E. Holzer, J. R. Spreiter, S. S. Stahara, and D. S. Chaussee, Solar wind flow about the terrestrial planets, 2, Comparison with gasdynamic theory and implications for solar-planetary interactions, J. Geophys. Res., **88**, 19, 1983a.
- Slavin, J. A., E. J. Smith, P. R. Gazis, and J. D. Mihalov, A pioneer voyager study of the solar wind interaction with Saturn, Geophys. Res. Lett., **10**, 9, 1983b.

- Slavin, J. A., R. E. Holzer, J. R. Spreiter, and S. S. Stahara, Planetary Mach cones: theory and observation, submitted to J. Geophys. Res., 1983c.
- Smith, E. J., R. H. Holzer, M. G. McLeod, and C. T. Russell, Magnetic noise in the magnetosheath in the frequency range 3-300 Hz, J. Geophys. Res., **72**, 4803-4814, 1967.
- Smtih, C. W., M. L. Goldstein, and M. B. Matthaeus, Turbulence analysis of the Jovian upstream 'wave' Phenomenon, J. Geophys. Res., **88**, 5581, 1983.
- Spreiter, J. R., Magnetohydrodynamic and gasdynamic aspects of the flow around the terrestrial planets: a critical review, solar wind interactions with the planets Mercury, Venus, and Mars, edited, N. F. Ness, NASA SP-397, pp. 135-150, 1976.
- Spreiter, J. R., and A. W. Rizzi. Aligned flow solutions for solar wind flow past the earth's magnetosphere, Acta. Astronaut., **1**, 15, 1974.
- Spreiter, J. R., and S. S. Stahara, A new predictive model for determining solar wind-terrestrial planet interactions, J. Geophys. Res., **85**, 6769, 1980.
- Spreiter, J. R. A. L. Summers, and A. Y. Alksne, Hydromagnetic flow around the magnetosphere, Planet. Space Sci., **14**, 223, 1966.
- Tanaka, M., C. C. Goodrich, D. Winske, K. Papadopoulos, A source of the backstreaming ion beams in the foreshock region, J. Geophys. Res., **88**, 3046, 1983.
- Tatrallyay, M., C. T. Russell, J. D. Mihalov, and A. Barnes, Factors controlling the location of the Venus bow shock, J. Geophys. Res., **88**, 5613, 1983.
- Thomsen, M. F., S. J. Schwartz, and J. T. Gosling, Observational evidence on the origin of ions upstream of the earth's bow shock, J. Geophys. Res., in press, 1983a.
- Thomsen, M. F., J. T. Gosling, S. J. Bame, W. C. Feldman, G. Paschmann, and N. Sckopke, Field-aligned ion beams upstream of the earth's bow shock: evidence for a magnetosheath source, Geophys. Res. Lett., submitted, 1983b.
- Thomsen, M. F., H. C. Barr, S. P. Gary, W. C. Feldman, and T. E. Cole, Stability of electron distributions within the earth's bow shock, J. Geophys. Res., **88**, 1983c.
- Tidman, D. A., and N. A. Krall, Shock Waves in Collisionless Plasmas, Wiley-Interscience, N.Y., 1971.
- Tokar, R. L., D. A. Gurnett, and W. C. Feldman, Whistler mode turbulence generated by electron beams in Earth's bow shock, J. Geophys. Res., in press, 1983.
- Tsurutani, B. T., E. J. Smith, R. R. Anderson, K. W. Ogilvie, J. D. Scudder, D. N. Baker, and S. J. Bame, Lion roars and nonoscillatory drift mirror waves in the magnetosheath, J. Geophys. Res., **87**, 6060, 1982a.
- Tsurutani, B.T., E. J. Smith, K. R. Pyle, and J. A. Simpson, Energetic protons accelerated at corotating shocks: Pioneer 10 and 11 observations from 1 to 6 AU, J. Geophys. Res., **87**, 7389, 1982b.

- Tsurutani, B. T., E. J. Smith, and D. E. Jones, Waves observed upstream of interplanetary shocks, J. Geophys. Res., in press, 1983.
- Vaisberg, O. L., A. A. Galeev, G. N. Zastenker, S. I. Klimov, M. N. Nozdrachev, R. Z. Sagdeev, A. Yu. Sokolov, and V. D. Shapiro, Electron acceleration in front of intense collisionless shock waves, Sov. Phys. JETP, in press, 1983.
- Wenzel, K. P., R. Reinhard, and T. R. Sanderson, Interplanetary acceleration of low-energy protons as observed during the 25 September 1978 shock event, Adv. Space Res., **1**, 3, 105, 1981.
- Wenzel, K. P., T. R. Sanderson, and P. VanNes, private communication, 1983.
- Wu, C. S., Physical mechanisms for turbulent dissipation in collisionless shock waves, Space Sci. Rev., **32**, 83, 1982.
- Wu, C. S., and R. W. Fredricks, Cyclotron drift instability in the bow shock, J. Geophys. Res., **77**, 5585, 1972.
- Wu, S. T., S. M. Han, and M. Dryer, Two-dimensional, time dependent MHD description of interplanetary disturbances: simulation of high speed solar wind interactions, Planet. Space Sci., **27**, 255, 1979.
- Wu, C. S., C. S. Lin, and H. K. Wong, Plasma waves excited in the upstream region by density gradient of energetic particle streams associated with the earth's bow shock, Geophys. Res. Lett., **8**, 393, 1981.
- Wu, C. C., R. J. Walker, and J. M. Dawson, A three-dimensional MHD model of the earth's magnetosphere, Geophys. Res. Lett., **8**, 523, 1981.
- Wu, C. S., Y. M. Zhou, S. T. Tsai, S. C. Guo, D. Winske, and K. Papadopoulos, A kinetic cross-field streaming instability, Phys. Fluids, **26**, 1259, 1983a.
- Wu, C. S., D. Winske, Y. M. Zhou, S. T. Tsai, P. Rodriguez, M. Tanaka, K. Papadopoulos, K. Akimoto, C. S. Lin, M. M. Leroy, and C. C. Goodrich, Microinstabilities associated with a high Mach-number, perpendicular shock, Space Sci. Rev., **34**, in press, 1983b.
- Zhou, Y. M., H. K. Wong, C. S. Wu, and D. Winske, Lower-hybrid-drift instability with temperature gradient in a perpendicular shock wave, J. Geophys. Res., **88**, 3026, 1983a.
- Zhou, Y. M., Y. Y. Li, and C. S. Wu, Stabilizing effect of magnetic field and temperature gradients on the electron cyclotron drift instability, submitted to Phys. Fluids, 1983b.
- Zwan, B. J., and R. A. Wolf, Depletion of solar wind plasma near a planetary boundary, J. Geophys. Res., **81**, 1636, 1976.
- Zwickl, R. D., S. M. Krimigis, J. F. Carbary, E. P. Keath, T. P. Armstrong, D. C. Hamilton, and G. Gloeckler, Energetic particle events ( $>30$  keV) of Jovan origin observed by Voyager 1 and 2 in interplanetary space, J. Geophys. Res., **86**, 796, 1980.
- Zwickl, R. D., S. M. Krimigis, J. F. Carbary, E. P. Keath, T. P. Armstrong, D. C. Hamilton, and G. Gloeckler, Energetic particle events ( $>30$  keV) of Jovian origin observed by Voyager 1 and 2 in interplanetary space, J. Geophys. Res., **86**, 8125, 1981.

## CHAPTER 11

# ASSESSMENT OF PLASMA TRANSPORT AND CONVECTION AT HIGH LATITUDES

### WORKING GROUP MEMBERS

Robert Schunk, Chairman  
*Utah State University*

A. R. Barakat  
*Utah State University*

Herbert Carlson  
*Air Force Geophysics Laboratory*

J. B. Evans  
*MIT Haystack Observatory*

John Foster  
*MIT Haystack Observatory*

R. Greenwald  
*Johns Hopkins University*

M. C. Kelley  
*Cornell University*

Thomas Potemra  
*Johns Hopkins University*

M. H. Rees  
*University of Alaska*

A. D. Richmond  
*National Oceanic and Atmospheric Administration*

R. G. Roble  
*National Center for Atmospheric Research*

**CHAPTER 11**  
**ASSESSMENT OF PLASMA TRANSPORT AND CONVECTION AT HIGH LATITUDES**

I.	Introduction	11-3
II.	Currents and Conductivities	11-4
III.	Thermospheric Winds and Plasma Transport	11-8
IV.	Plasma Convection	11-13
V.	Polar Wind	11-27
VI.	Conclusions	11-31
VII.	References	11-32

## I. INTRODUCTION

The high-latitude ionosphere is strongly coupled to the thermosphere and magnetosphere, as shown schematically in Figure 11-1. The magnetospheric coupling occurs via electric fields, field-aligned currents, and particle precipitation. Owing to the interaction of the shocked solar wind with the geomagnetic field, an electric potential difference is generated across the tail of the magnetosphere, with the resulting electric field pointing from dawn to dusk. Except for isolated regions, typically in the auroral oval, the geomagnetic field lines are equipotentials due to the high electrical conductivity along field lines. Consequently, this cross-tail potential difference is mapped into the high-latitude ionosphere as an electric field that is directed perpendicular to the geomagnetic field. At ionospheric heights, this perpendicular (or convection) electric field is typically  $25\text{--}50\text{ mV m}^{-1}$  in the polar cap, but can be much greater than  $100\text{ mV m}^{-1}$  in restricted latitudinal bands at certain times.

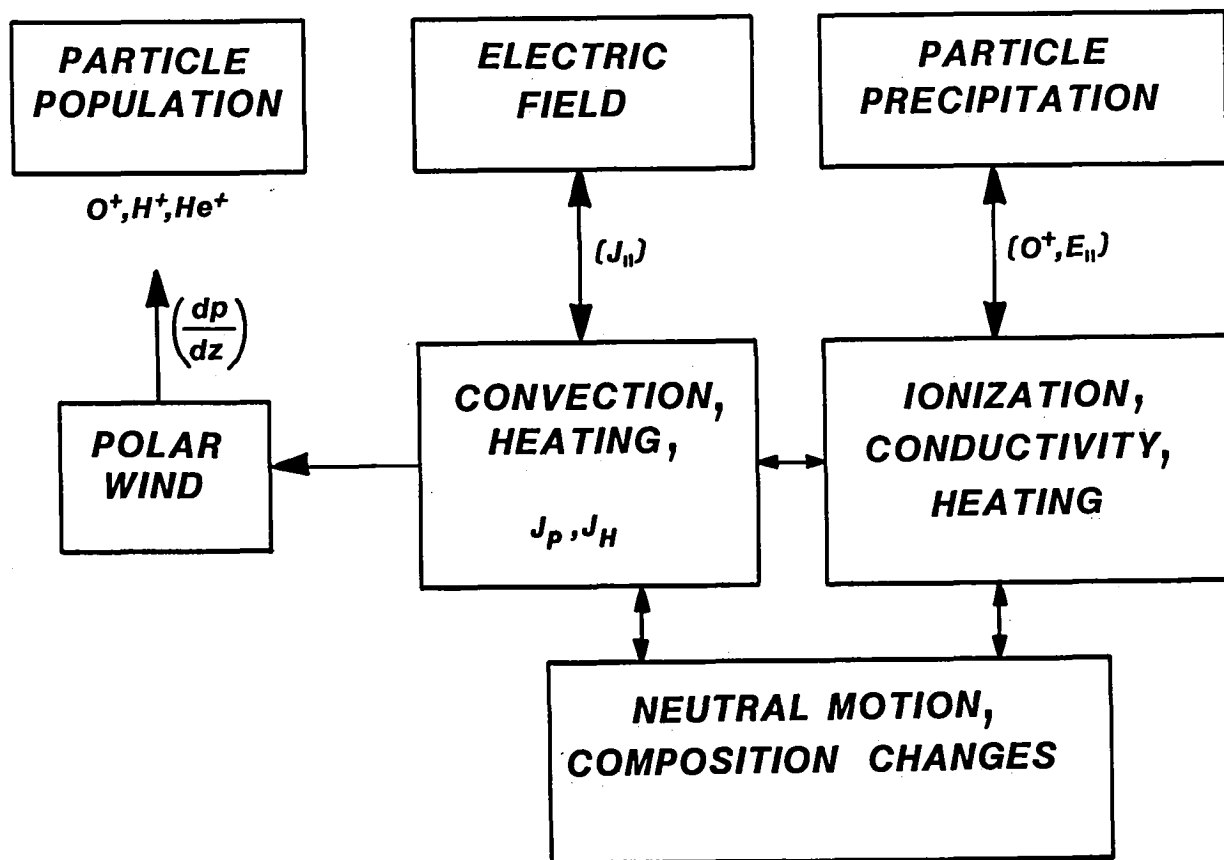


Figure 11-1. Block diagram showing the coupling between the thermosphere, ionosphere, and magnetosphere.

The effect that the convection electric field has on the ionosphere depends on altitude. At F-region altitudes and above, both ions and electrons drift in the  $\mathbf{E} \times \mathbf{B}$  direction, where  $\mathbf{E}$  is the convection electric field and  $\mathbf{B}$  is the geomagnetic field. The convection pattern is basically a two-cell pattern with antisunward flow over the polar cap and return flow at lower latitudes. At low altitudes (E region), the ion velocity vector rotates toward the  $\mathbf{E}$  direction owing to the increased ion-neutral collision frequency, while the electrons continue to drift in the  $\mathbf{E} \times \mathbf{B}$  direction. Consequently, large-scale horizontal currents ( $J_p$  and  $J_H$ ) flow in the E region. These horizontal ionospheric currents are coupled to the magnetosphere through field-aligned (Birkeland) currents. The magnitude and direction of the currents as well as the heating associated with current dissipation depend not only on the magnetospheric generator, but also on the neutral gas motion. Likewise, the neutral gas is strongly affected by the convecting plasma through direct momentum forcing and frictional heating, which act to induce changes in the thermospheric wind and composition. These changes, in turn, affect the ionospheric densities, conductivities and currents.

Energetic particle precipitation from the magnetosphere in the auroral region leads to the creation of ionization and to electron, ion, and neutral gas heating. These processes, in turn, affect the ionospheric conductivities and currents as well as the neutral wind and composition. The precipitating electrons gain energy from electric fields that are directed parallel to the geomagnetic field, which occur only in the auroral region. The parallel electric fields also produce energetic, upstreaming  $O^+$  ions ( $\sim keV$ ), which are an important ionospheric source for the magnetosphere.

The ionosphere also is a source of both thermal ( $\sim 0.5$  eV) and suprathermal ( $\sim 5$  eV) ions via the polar wind. At altitudes greater than about 800 km, the ionospheric ions can escape from the top-side ionosphere along geomagnetic field lines owing to the pressure difference ( $dp/dz$ ) between the ionosphere and magnetosphere; a process termed the polar wind. Substantial fluxes of  $H^+$ ,  $He^+$  and  $O^+$  ions continually flow into the magnetosphere on geomagnetic field lines greater than about  $60^\circ$  (invariant latitude).

In order to assess our current understanding of plasma transport and convection at high latitudes, it is necessary to take account of the strong coupling between the ionosphere, thermosphere, and magnetosphere. This will be done in the following subsections.

## II. CURRENTS AND CONDUCTIVITIES

Birkeland currents link the interplanetary medium and the near-Earth environment. They convey changes in the interplanetary medium, such as the orientation of the IMF to the Earth's polar cap and auroral zone. These currents are a critical element in a variety of plasma processes associated with aurora, ionospheric scintillations, surface micropulsations and radio emissions.

The Birkeland currents flow into and away from the auroral zone along geomagnetic field lines forming "cone-shaped" regions, as shown in Figure 11-2. These currents are a permanent feature of the high latitude regions and they flow in specific directions. On the morning side, they flow downward on the polar side, referred to as "region 1" currents, and they flow upward on the equator side, referred to as "region 2" currents. The flow pattern on the afternoon side is systematically reversed, i.e., upward flow on the polar side (region 1) and downward flow on the equatorward side (region 2).



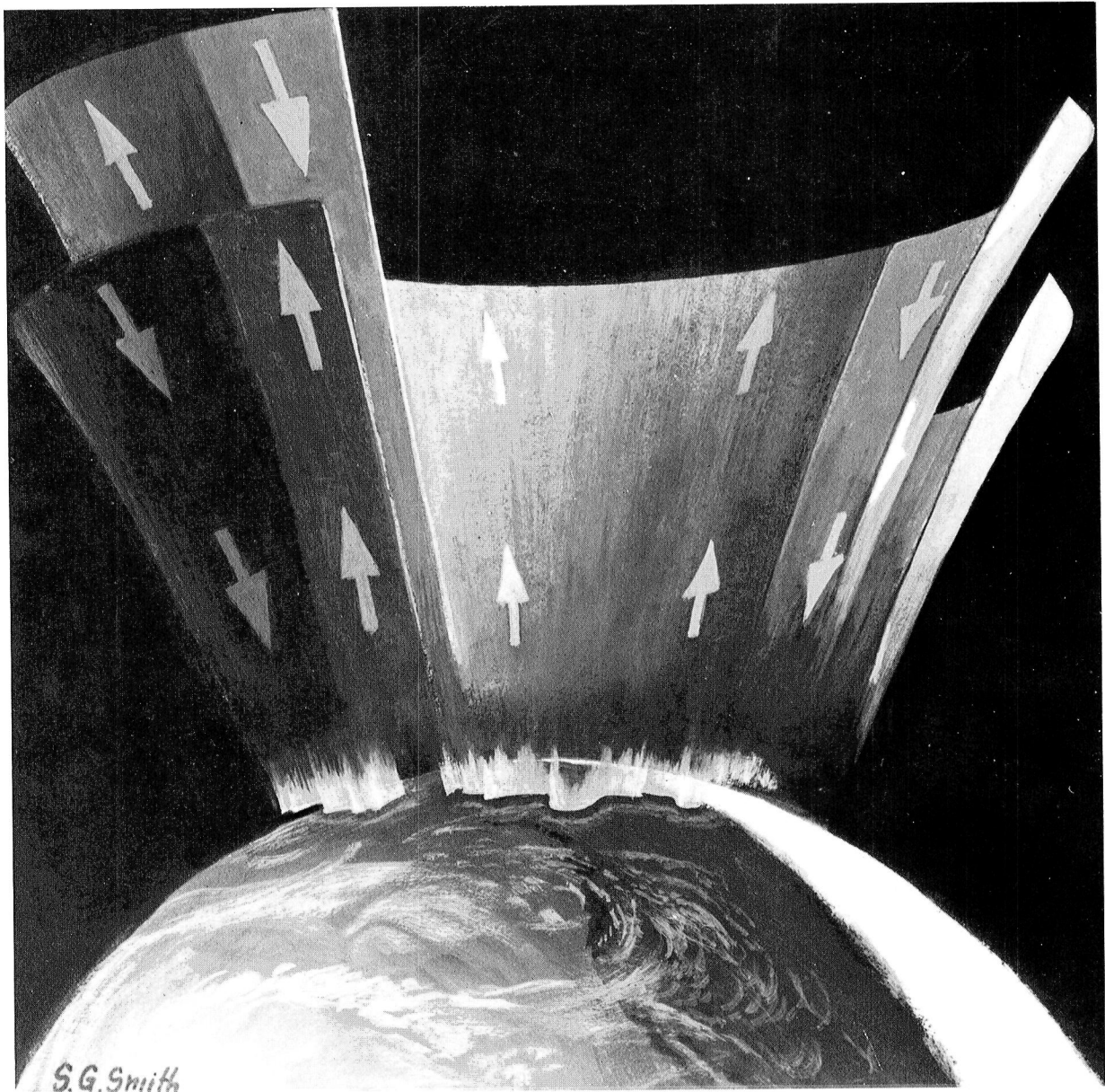


Figure 11-2. Cone-shaped regions formed by Birkeland currents flowing into and away from the auroral region. The currents on a given sheet reverse direction near midnight and noon.

The electric field that convects ionospheric plasma is intimately connected with electric currents flowing horizontally in the ionosphere and along geomagnetic field lines between the ionosphere and outer magnetosphere. The patterns and magnitudes of electric fields and currents are determined by the characteristics of the entire magnetospheric-ionospheric electrical circuit. In the ionosphere, Ohm's law relates the current density  $\underline{J}$  to the electric field in the frame of reference of the moving medium,  $\underline{E} + \underline{v} \times \underline{B}$ , ( $\underline{E}$  and  $\underline{B}$  being the electric and geomagnetic fields and  $\underline{v}$  the neutral wind velocity) by

$$\underline{J} = \sigma_{\parallel} \underline{E}_{\parallel} + \sigma_P (\underline{E}_{\perp} + \underline{v} \times \underline{B}) + \sigma_H \frac{\underline{B}}{B} \times (\underline{E} + \underline{v} \times \underline{B})$$

where  $\sigma_{\parallel}$ ,  $\sigma_P$ ,  $\sigma_H$  are the parallel, Pedersen and Hall conductivities, and subscripts " $\parallel$ " and " $\perp$ " denote vector components parallel and perpendicular to  $\underline{B}$ . The parallel conductivity is so large that  $E_{\parallel}$  must be very small, leading to the result that  $\underline{E}_{\perp}$  is nearly constant along magnetic field lines in the ionosphere. We often deal with height-integrated conductivities and current densities and treat the ionosphere as a conducting sheet. The magnitudes of the  $\underline{v} \times \underline{B}$  and  $\underline{E}$  terms are highly variable. The rate of transfer of electromagnetic energy to the weakly ionized gas is given by  $\underline{J} \cdot \underline{E}$ , most of which goes into heating of the gas, but some of which goes into acceleration. It is essential to examine the entire electrodynamic system, including both electric fields and currents, in order to understand high latitude convection and its effects on the upper atmosphere.

The strongest ionospheric currents tend to occur in the auroral oval, associated both with the strong electric field and with conductivity enhancements produced by precipitating electrons and ions from the magnetosphere. The ionospheric currents are coupled with field-aligned currents that are most intense around the auroral oval. The entire system of electric fields, currents, and conductivities is highly variable, which makes it difficult to obtain a clear picture of instantaneous spatial patterns and interrelationships among patterns of the fields, currents, and conductivities.

## MEASUREMENT TECHNIQUES

A large variety of techniques exist for measuring ionospheric electrodynamic parameters, each technique having peculiar strengths and limitations. Ground magnetometers are spread over most of the Northern Hemisphere and can determine the patterns of magnetic perturbations on a continual basis. Incoherent scatter radars can measure electron densities and ion drift velocities, giving conductivities, electric fields, and electric currents, with varying degrees of accuracy within their field of view. Other types of radio wave soundings also yield information on electron densities and drift velocities. The Doppler shift of emissions from neutral and ionic species yields measurements of the neutral wind and ion drifts, respectively. Polar orbiting satellites can measure magnetic perturbations, electric fields, particle precipitation, and optical emissions from aurora, yielding information about field-aligned currents, conductivities, and sometimes ionospheric currents. Most of the information is limited to the line of the satellite orbit, except for optical imaging in visible, UV, and X-ray wavelengths, which can yield the conductance over the entire field of view.

## THEORETICAL APPROACHES

The use of Ohm's law can provide information about unmeasured electrodynamic parameters when others are known. Theoretical modeling can also be used to test the consistency among independent measurements of the various parameters. Global modeling studies have been pursued from different directions. Starting from field-aligned current distributions and a conductivity model, the ionospheric electric field and current as well as magnetic perturbations can be calculated. The three-dimensional current is constrained to be divergence free and the electric

field is normally assumed to be electrostatic so that it can be represented by a scalar potential. Winds, if known, can be included in the calculations. A different type of approach is to start from an "equivalent current" distribution determined from ground magnetic data, together with a conductivity model, to calculate the electric fields and the ionospheric and field-aligned currents (Kamide et al., 1981). Another method to calculate current distributions from magnetic perturbation measurements is to fit physically realistic elemental current loops, including field-aligned currents, to the observations. This approach does not explicitly use Ohm's law, which has the advantage of not requiring a conductance model but the disadvantage of requiring independent information about the ratio of north-south to east-west ionospheric currents (Kisabeth, 1979).

## CONDUCTIVITIES

The ionospheric conductivity is proportional to the ionization density and is a function of the charged particle collision frequencies with neutral molecules, especially the ion-neutral collision frequency. The collision frequencies depend weakly on the ion species but vary strongly with altitude, as they are essentially proportional to the neutral density. Most determinations of ionospheric conductivities rely on a model neutral atmosphere together with measured or computed electron densities, although some attempts have been made to infer conductivities by comparing measured currents and electric fields.

XUV solar radiation is the principal source of ionization and hence of conductivity in the sunlit hemisphere, and its variability is normally slow (except during solar flares) so that sunlit conductivities can be reasonably well determined. In contrast, dark hemisphere conductivities are quite variable and much more difficult to determine accurately. In the auroral oval, in particular, conductivity enhancements from precipitating energetic particles play a dominant role in ionospheric electrodynamics. Statistically averaged patterns of auroral conductivities have been produced, parameterized by the magnetic activity indices  $K_p$  and AE. These indices cannot adequately represent the instantaneous pattern of conductivity enhancement, however. For example, there is evidence that the pattern of enhancement is different during the expansive and recovery phases of a substorm even for equal values of AE. More detailed models are therefore desirable. Developments are in progress to determine instantaneous conductivity enhancements over part or all of the auroral oval from satellite images of the aurora in visible, ultraviolet, and X-ray wavelengths.

## CURRENTS

The large-scale features of high-latitude currents and convective flow are consistent with the observed electric fields. For example, the general sunward plasma flow in the auroral zone produces electric fields directed equatorward in the dawn side (consistent with downward flowing Birkeland currents on the poleward side and upward flowing currents on the equatorward side) and directed poleward on the dusk side (associated with a similar consistent Birkeland current pattern). The system of eastward and westward auroral "electrojet" currents is also generally consistent with the convective flow and Birkeland current patterns. These "consistent relationships" apply mainly to the large-scale (>100 km) statistical (time scales limited by the 100 minute satellite orbital periods) features (Potemra, 1979).

Instantaneous data are more difficult to interpret, but some characteristic features of dynamic events have been identified, such as the nighttime westward auroral electrojet associated with substorms. Regions where the current configuration and its relations to the electric field, visible aurora, and field-aligned currents have not been fully delineated are in the dayside cusp, around the nighttime Harang discontinuity, and near the boundary of the auroral oval with the polar cap. These regions are particularly important because they appear to map to boundary regions in the magnetosphere associated with significant energy transfer processes.

For phenomena as dynamic and variable as high-latitude electric currents a combination of observational and theoretical studies is required to define sufficient constraints on allowed patterns and interrelationships so that understanding of the system can be achieved. Because of the intimate coupling with magnetospheric processes, this understanding will result only from a comprehensive examination of the entire magnetosphere-ionosphere system.

## MAJOR UNRESOLVED PROBLEMS

- What is the relationship between Birkeland currents and ionospheric currents and electric fields? How do they vary in time?
- What is the role that Birkeland currents play in ionospheric plasma instabilities (such as micropulsations, AKR, ion beams and conics, acceleration processes, and ionospheric scintillations)?

## DIRECTION OF FUTURE RESEARCH

We need a global conductance pattern to connect the ionospheric currents to the magnetospheric system of currents, convective flow, and electric fields. Without this global conductance pattern, the surface measurements are, in effect, shielded from the entire space above the ionosphere.

An important direction for theoretical work on high-latitude currents is the development and improvement of techniques designed to incorporate simultaneous measurements of electric fields, magnetic perturbations, and conductances into a coherent, self-consistent, time-dependent, electrodynamic model of dynamic events.

The adequacy and limitations of the theoretical conductivity expressions commonly used need to be investigated. It is not clear how theoretical expressions derived for ion drift velocities, small in comparison with the speed of sound, apply when drifts are comparable to or exceed the sound speed. Plasma instabilities occur when the electric field is large, but their quantitative influence on the effective conductivities is not known. The presence of small-scale spatial structure in the auroral zone conductivities probably affects the relation between large-scale currents and electric fields, essentially modifying the mean conductivity. The evaluation of this effect will require careful theoretical and observational analysis of the small-scale spatial relations between conductivities, currents, and electric fields.

## III. THERMOSPHERIC WINDS AND PLASMA TRANSPORT

### IONOSPHERE - THERMOSPHERE COUPLING

The ion drift velocity appears in the equation of motion for the neutral gas, and the neutral gas velocity appears in the equation of motion for the ions. The coupling term is the ion drag,  $\nu_{in}(\underline{U} - \underline{V})$ , where the collision frequency for momentum transfer is  $\nu_{in}$ , and  $\underline{U}$  and  $\underline{V}$  are the neutral air and ion velocities, respectively. The ion-neutral,  $\nu_{in}$  and the neutral-ion  $\nu_{ni}$  collision frequencies are related by Newton's third law,  $\rho_n \nu_{ni} = \rho_i \nu_{in}$ , where the  $\rho$ 's are the densities of the neutrals and ions. In the daytime F region the ion to neutral abundance ratio is about 1:1000,  $\nu_{in} \approx 1 \text{ sec}^{-1}$  and  $\nu_{ni} \approx 10^{-3} \text{ sec}^{-1}$ . Thus, the time constant associated with ions communicating their motion (as induced by an electric field) to the neutrals is about 1000 sec. However, the time constant associated with a neutral wind communicating its motion to the ions is

only one second. The plasma responds promptly to a neutral wind. The components of the collisionally-driven plasma drift are given by  $\underline{U} \cdot \underline{B}$  and  $\underline{U} \times \underline{B}$ .

The ambipolar diffusion velocity,  $V_a$ , is a function of the neutral gas velocity  $U \sin I \cos I$ , where  $I$  is the dip angle relative to the horizontal. The magnitude of the neutral wind at high latitudes can be large and the effect cannot be ignored in computing plasma diffusion even though the dip angle dependence decreases the field-aligned components of a horizontal wind.

There is also energy exchange in collisions between neutrals and ions which is proportional to the sum of a velocity term,  $v_{in}(\underline{U} - \underline{V})^2$ , and a term that describes the temperature difference between two gases,  $v_{in}(T_i - T_n)$ .

It is probably fair to note that current research in neutral gas-plasma interaction emphasizes the effect of ion motion upon the neutral wind, the ion motion being the specified input parameter, and we describe the current understanding of this problem next.

High-latitude plasma convection has an important influence on the global thermospheric structure and circulation (Figure 11-3); (Roble et al., 1982; Rees et al., 1980; Smith et al., 1982; Rishbeth, 1972). Recent modeling has shown that the momentum source associated with ion convection drives a largely rotational, nondivergent, double-vortex neutral wind system at F-region altitudes that can attain velocities greater than  $500 \text{ ms}^{-1}$  during moderate levels of geomagnetic activity. These winds follow but lag the ion drift pattern by a few hours and they have been observed by the Dynamics Explorer-2 satellite. At F-region heights the ion convection velocity is almost divergence-free, so that vertical motions are small and convection-driven neutral thermospheric winds are not seriously modified by 'back pressure'. Therefore, F-region neutral winds at high latitudes tend to follow the ion drift pattern. Joule heating associated with plasma convection, on the other hand, drives vertical motions that act to reduce the pressure and temperature gradients due to heating, and hence, the horizontal winds. Previous studies have shown that the various components of the neutral gas react to these two forcings in a different manner. It is primarily the vertical motions due to Joule heating that cause the observed depletion of atomic oxygen and helium and the observed enhancements of  $\text{O}_2$ ,  $\text{N}_2$ , and Ar in the higher latitude regions (Hays et al., 1973). Plasma drift-driven winds, however, produce only small deviations from diffusive equilibrium distributions even though they are accelerated to high velocities at F-region heights.

In the lower thermosphere the ion drift pattern rotates further from the  $\underline{E} \times \underline{B}$  drift direction due to the change in the Pedersen and Hall conductivities with height. As a result, the double-vortex wind system at F-region heights that extends downward into the lower thermosphere is modified by nonlinear processes. The cold low-pressure cyclonic circulation near the dawn terminator is much more pronounced than the warm high pressure anticyclonic circulation in the evening sector. The dawn-to-dusk temperature contrast across the polar cap and the intensity of the cyclonic and anti-cyclonic circulation cells are both related to the intensity of the cross-tail potential that drives magnetospheric plasma convection. Since Joule heating is related to the relative drift between the ions and neutrals it is important to calculate the neutral gas motion self-consistently to determine the plasma heating rates. The largest Joule heating rates are expected near the boundary between the antisunward plasma flow over the polar cap and the sunward plasma flow near the auroral zone. The inertia of the neutral gas at times results in oppositely directed ion drift and neutral winds, especially in the dawn sector where the largest Joule heating rates have been observed by both incoherent scatter radars and the DE-2 satellite.

In both the upper and lower thermosphere, plasma convection with displaced poles introduces a Universal Time dependence to the circulation, temperature and compositional structure; similar patterns occur in the Northern and Southern hemisphere, with a 12-hour time difference (Figure 11-3).

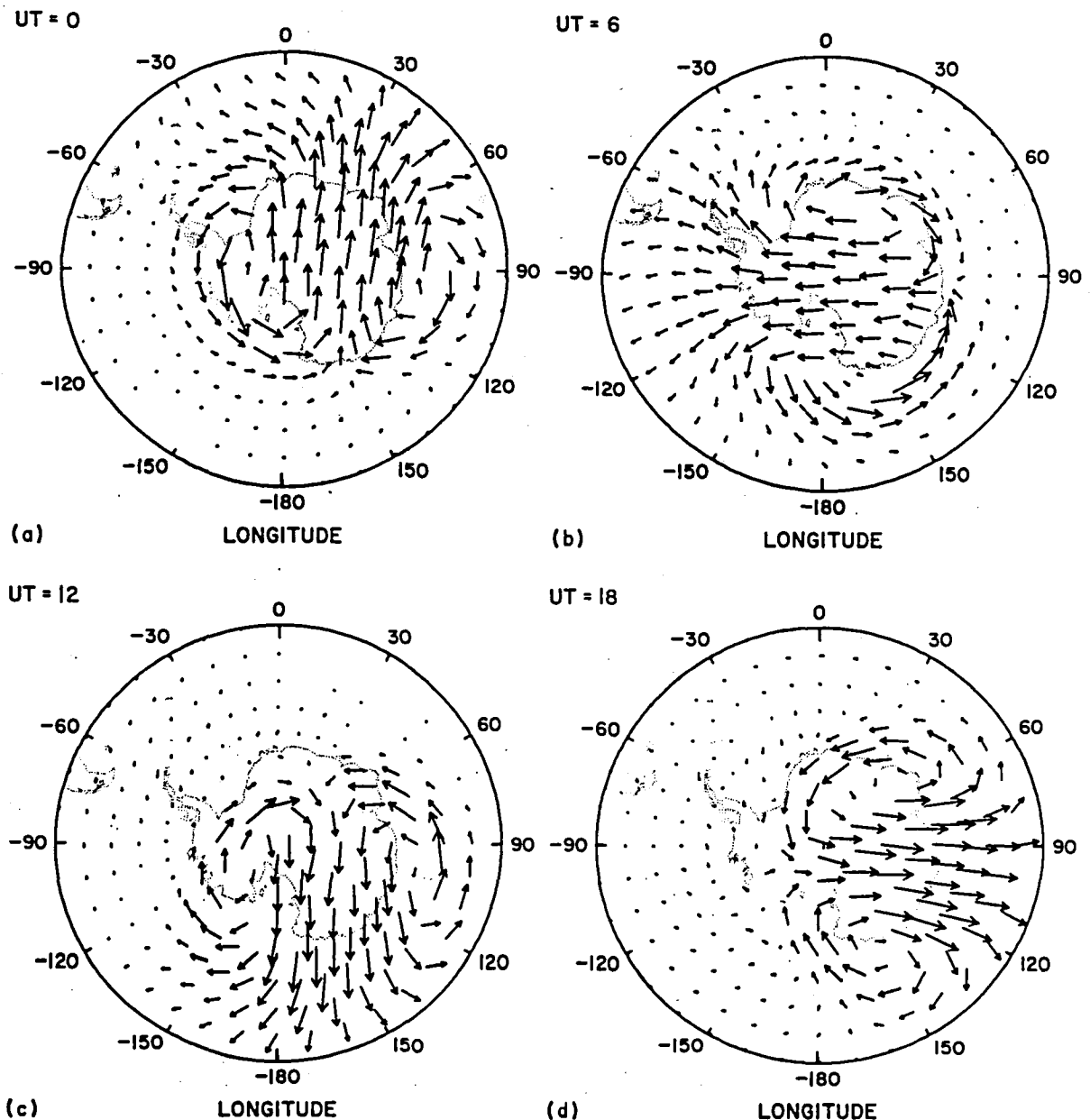


Figure 11-3. Polar plots of the calculated circulation over the Southern Hemisphere (summer) polar region during late October for the case of solar heating plus magnetospheric convection with a cross-tail potential of 60 kV. The geomagnetic and geographic poles are displaced and the circulation at four UT's are given: (a) 00 UT, (b) 06 UT, (c) 12 UT, and (d) 18 UT. The maximum arrow represents (a)  $567 \text{ ms}^{-1}$ , (b)  $514 \text{ ms}^{-1}$ , (c)  $391 \text{ ms}^{-1}$ , and (d)  $400 \text{ ms}^{-1}$ . From Roble et al. (1982).

During geomagnetic storms and substorms, plasma convection is enhanced considerably. The winds at F-region heights respond to the enhanced ion drag by accelerating and decelerating with changes in ion drift velocity on time scales of 1/2-2 hours depending upon the magnitude of the electron density. In the lower thermosphere intense vortices are generated by plasma convection and persist long after the storm or substorm subsides. These vortices spin down slowly due to the inertia of the neutral gas and a relatively small ion drag damping, as compared to F-region altitudes. As a result, these E-region vortices can produce, by dynamo action, electric fields directed radially inward or outward from the vortex center depending upon whether the vortex is clockwise or counterclockwise and thus produce field-aligned currents that flow to or from the magnetosphere with a duration that is comparable to the lifetime of the neutral gas vortex. This dynamo generated electric field and current may interact with the magnetospheric plasma convection pattern in a complex manner. Thus, it is important to consider the plasma drift at high latitudes in the context of a coupled system with mutual interactions between the plasma and dynamics of the neutral atmosphere.

Auroral particle precipitation influences plasma drift at high latitudes through the resulting ionization enhancement that governs the Pedersen and Hall conductivities. Not only are the magnitudes of the conductivities of interest, but also the altitude profiles of the two orthogonal components since these determine the direction of the ionospheric current. In the course of energy degradation of the particle beam a small fraction of the energy goes into heating the ions and the ambient electrons, resulting in an enhancement of the plasma temperatures,  $T_e$  and  $T_i$ , parameters that appear in the plasma transport governing equations (Figure 11-4).

The strongest influence of energetic particle precipitation on plasma transport is through neutral gas heating (Figure 11-5); (Rees et al., 1983). More than half the energy deposited in the atmosphere goes into neutral heating, and this localized heating is a driver for the high-latitude thermosphere.

## COMPARISON OF THEORY AND OBSERVATIONS

Current theory has focused on the effects of plasma motion on neutral thermospheric dynamics and, to our knowledge, nothing has been done on the inverse problem on a global scale.

Simultaneous neutral wind and ion drift measurements were obtained at Chatanika, Alaska and are currently carried on at Sondrestrom, Greenland. These are data obtained with optical interferometers measuring the doppler shift of the 6300Å radiation of atomic oxygen, and the incoherent scatter radar measuring the ion drift velocity.

There has been work done on the effects of neutral wind shears producing thin layers of ionization in the E region and enhanced ionospheric currents. Observations and theory have, however, only been compared at mid-latitudes where the magnetic dip angle is smaller than at high latitudes (Rishbeth and Walker, 1982).

## MAJOR UNRESOLVED PROBLEMS

Neutral gas motion and plasma transport are coupled, but under what conditions is the neutral wind the driver or the brake for plasma motion? Once convection electric fields set the high latitude neutral gas into motion, the neutral atmosphere will continue to drive a nominal convection pattern even if the magnetospheric driving force is removed. What effect will this have on magnetospheric transport, or for that matter, magnetospheric substorms? The extent to which the atmospheric dynamo affects magnetospheric processes needs to be assessed, including diurnal, latitudinal, seasonal, and geomagnetic activity influences.

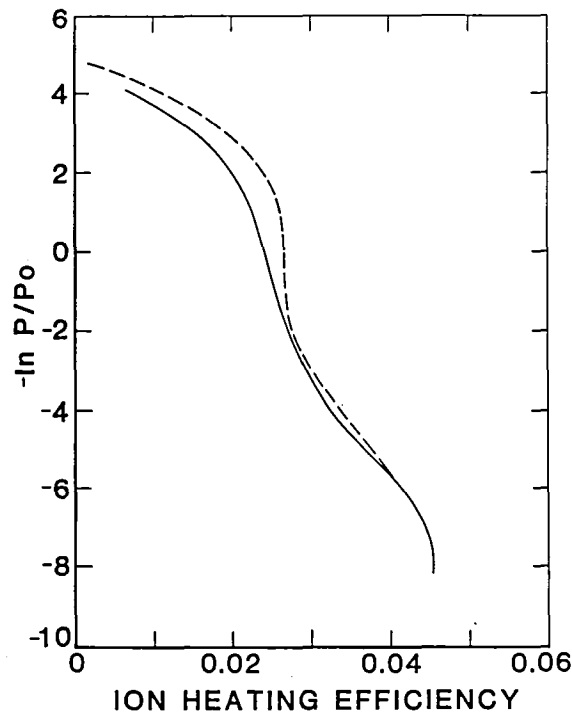


Figure 11-4. Ion heating efficiency in log pressure coordinates for solar maximum (solid curve) and solar minimum (dashed curve) conditions at 65°N latitude in December at midnight. From Rees et al. (1983).

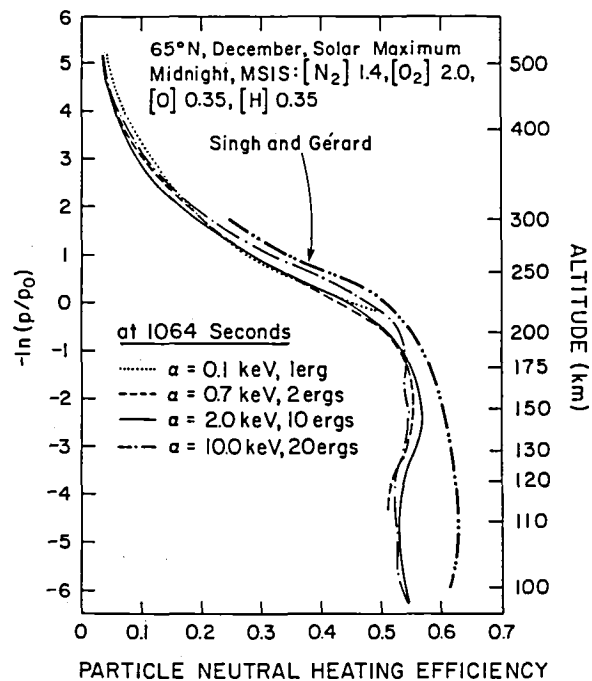


Figure 11-5. Generalized log pressure profiles of the neutral gas heating efficiency for auroral electron fluxes with Maxwellian energy spectra of several characteristic energies. The results obtained by Singh and Gerard (1982) for a similar neutral atmosphere are also shown. From Rees et al. (1983).



Finally, there is the question of the degree to which large scale wave phenomena affect thermospheric dynamics. Energy may be transported into the thermosphere by gravity waves propagating upwards from the troposphere and stratosphere. Also, large scale MHD resonance modes may transport energy from the magnetosphere into the high latitude ionosphere via A.C. Joule heating. Preliminary work on this effect has been done by Greenwald and Walker (1980).

## DIRECTION OF FUTURE RESEARCH

It would be desirable to obtain a body of simultaneous measurements of neutral wind and ion drift along the auroral oval (including the cusp region) and in the polar cap. Simultaneous data are required from all locations in order to separate local time and universal time effects, and the frequency of the measurements should be commensurate with changes in substorm phenomena and response times of ion-neutral interactions.

Modeling efforts should be directed toward self-consistent solutions of the plasma and neutral gas dynamical equations in three dimensions and in a time-dependent formulation. Sensitivity tests should be applied systematically to parameters that characterize the neutral motion as well as the plasma drift. The boundary and initial conditions should be guided by the observations obtained from the experimental program described above.

## IV. PLASMA CONVECTION

The high latitude plasma convection and associated electric fields have been studied by a variety of techniques including satellite-borne and rocket-borne double-probe and plasma drift measurements (Heppner, 1972a, b, 1977; Maynard, 1974; Heelis et al., 1976; Kelley et al., 1971, 1975), balloon-borne probe measurements (Mozer and Lucht, 1974), observations of the drift of clouds of ionized barium (Haerendel et al., 1967), incoherent scatter observations of the drift of the F-region plasma (Doupnik et al., 1972; Banks and Doupnik, 1975), coherent scatter observations of the drift of E-region irregularities (Greenwald et al., 1978; Baumjohann et al., 1980), and optical interferometer measurements of drifting  $O^+$  ions. The significance of these results to an overall understanding of the interaction of the solar wind with the Earth's magnetosphere has been reviewed by Banks (1975), Stern (1977), and Greenwald (1983), among others.

All of the measurements reported to date have provided information on only a limited spatial region at any instant of time, requiring that the overall pattern of high latitude electric fields be synthesized by combining observations made at a variety of places and times. In the case of studies with polar-orbiting satellites, this synthesis is of a number of passes over a significant portion of the high-latitude convection pattern. Each pass covers an extremely small longitudinal section of the pattern. For balloons and barium releases, the measurements are made at a single point and synthesis is provided by a large number of different observations and by conducting the measurements over an extended period of time (hours to days). The radar and optical measurements, depending on the technique and facility, are generally restricted to  $4^\circ$ - $17^\circ$  in latitude and a similar extent in longitude. Often the local time dependence of the radar measurements is obtained by conducting measurements for periods ranging from 1-2 days.

Besides these direct observations of convection flow, efforts have been made to infer the pattern indirectly from: (a) averaged magnetic perturbations at the surface of the Earth or (b) field-aligned averages of currents flowing into the ionosphere from the magnetosphere (Potemra 1979). In either case, a model for the ionospheric conductivity must be adopted. Work along

these lines has been published by Nisbet et al. (1978), Kamide and Matsushita (1979a, b) and others. See subsection on currents and conductivities where this is discussed in more detail.

## AVERAGE CONVECTION MODELS

While individual measurements show the convection pattern to be both structured and highly variable, responding to substorms as well as to changes in the IMF and ionospheric conductivity, there appears to be an underlying pattern which can be approximated by a double cell structure. This pattern serves as a useful basis for modeling the coupling between high-latitude convection and the neutral atmosphere as well as for studies of magnetosphere-ionosphere coupling. To date, our understanding of the dynamic behavior of this pattern has been limited to its dependence on directional changes in the IMF and  $K_p$ .

Heppner (1977) was the first to attempt a quantitative model for the convection pattern at ionospheric heights. He published two models for medium levels of geomagnetic activity ( $K_p \sim 3$ ), corresponding to the IMF directed "towards" and "away" from the Sun. These models were based on a limited set of OGO-6 single-axis electric field measurements along the dawn-dusk meridian together with the results of barium releases in other time sectors. Despite these limitations, these models were consistent with the drifts observed from incoherent scatter radars (Evans et al., 1980). An example of a comparison between Heppner's results and observations from Millstone Hill for a "toward" IMF configuration is shown in Figure 11-6.

High-latitude drift observations by the Atmospheric Explorer satellites cover nearly the complete range of local times. An empirical electric field model, including details of the convection pattern in the complex noon and midnight local time sectors, has been presented by Heelis et al. (1982). Ion drift velocities observed along a single orbit and a representation of the AE model are shown in Figure 11-7.

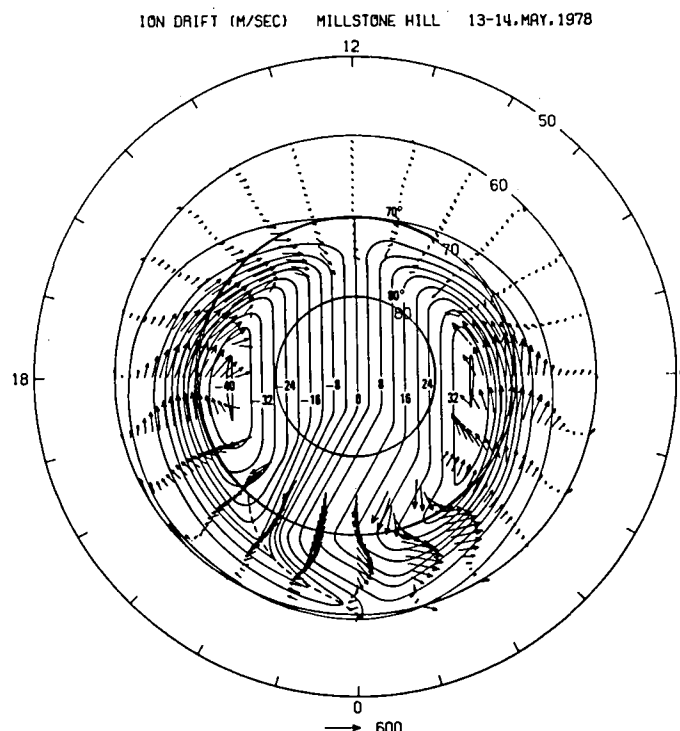


Figure 11-6 Superposition of Millstone Hill drift results for 13-14 May 1978 and Heppner's Model "A" for the case of the IMF towards the Sun. From Evans et al. (1980)

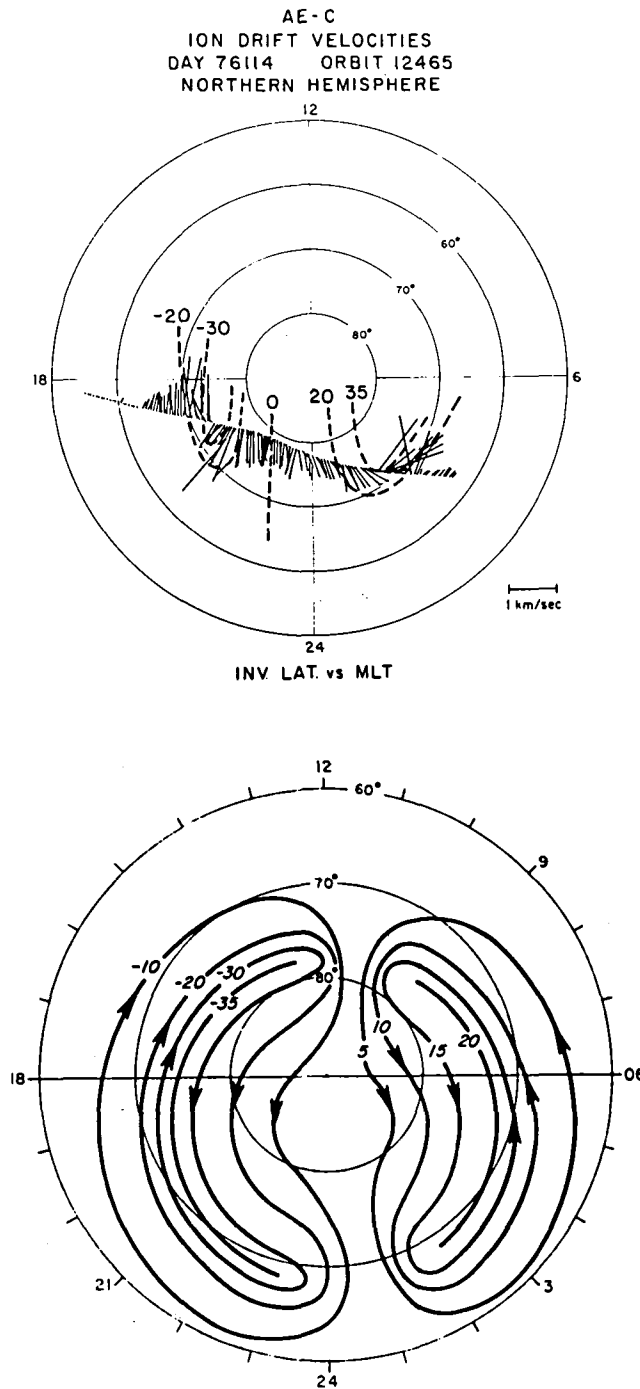


Figure 11-7. (a) Ion drift velocities measured by the Atmosphere Explorer satellite (AE-C) as it traversed the northern high-latitude ionosphere. Note the existence of both shear and rotational convection reversals. From Heelis and Hanson (1980). (b) Empirical convection model based on Atmosphere Explorer satellite data. The model takes account of convection features such as the dayside throat and the Harang discontinuity. From Heelis et al. (1982).

Several models of the average pattern of high latitude convection electric fields have been derived from incoherent scatter radar observations. Foster (1983) has presented an empirical model spanning the latitude range  $58^\circ\Lambda$ - $75^\circ\Lambda$  with  $1/2^\circ$  resolution based on Chatanika incoherent scatter radar observations during nine summer days for moderately disturbed conditions ( $K_p=3$ ). A cross polar cap potential of approximately 70 kV is represented. Oliver et al. (1983) have used all of the observations from Millstone Hill in 1978, gathered during 19 runs of greater than 24 hours and spanning the latitude range  $60^\circ\Lambda$ - $75^\circ\Lambda$ , to construct models for three different levels of  $K_p$ . The seasonal and IMF dependencies of the convection pattern are also being investigated with the radar data base. Thus far, no radar data have been collected at latitudes greater than  $75^\circ\Lambda$ , and the extension of radar deduced patterns to polar cap latitudes is uncertain.

Based on the currently available measurements, the dominant character of the convection pattern, as viewed in a corotating reference frame, can be represented by two cells centered approximately on dusk and dawn. On the basis of limited data, there are indications that under extremely quiet conditions ( $B_z$  positive), the antisunward flow of plasma across the polar cap may be replaced by two additional cells with flow paths that are reversed from those in the dawn and dusk sectors. This condition has not yet been modeled empirically.

With reference to the average convection patterns, the following conclusions can be drawn:

- There is an expansion and intensification of the convection pattern with  $K_p$  for IMF towards or away from the Sun.
- For both IMF towards and away from the Sun, the location of plasma entry into the dayside polar cap is in the pre-noon sector. For an away IMF, the entry point appears to move to earlier local times with increasing  $K_p$ .
- For both IMF cases, the dusk convection cell is generally larger than the dawn cell.
- The Harang discontinuity, which can be clearly seen on individual days, is often smoothed over in the averages. This suggests that its location is highly variable.
- For the IMF towards the Sun, antisunward (i.e., polar cap flow), can be seen in the dusk sector at  $\Lambda > 70^\circ$  (but not in the dawn sector) for large  $K_p$ .
- For the IMF away from the Sun, antisunward flow can be seen in the dawn sector at  $\Lambda > 70^\circ$  (but not in the dusk sector), for large  $K_p$ .

These last two observations suggest that the width of the dusk cell is reduced when the IMF is towards the Sun and may be kidney-shaped, possibly wrapped around the dawn cell, which then is oval-shaped. When the IMF is away, the dawn cell may become kidney-shaped and be wrapped around a larger oval dusk cell.

## TWO-DIMENSIONAL STRUCTURE OF CONVECTION

While the large-scale convection pattern at auroral latitudes is adequately represented by the currently available models, individual observations reveal that many of the features of the pattern are more sharply defined than the models suggest. In addition, there can exist significant fine scale structure associated with auroral activity, the location of auroral precipitation features, and the immediate pre-history of the ionospheric plasma. Observational techniques have been developed to display such detailed features in order to assess their characteristics and variability.

By scanning the Millstone Hill radar antenna in azimuth, the convection pattern is observed over  $25^\circ$  of latitude and 5 hours of local time in less than 10 minutes. Figure 11-8 presents a representative evening sector potential pattern, which reveals a 40 kV potential variation across auroral latitudes and a rotational reversal at the auroral oval/polar cap boundary above  $75^\circ\Lambda$ . In the case shown, the convection potential does not extend below  $60^\circ\Lambda$ .

Figure 11-9 exhibits an example of the plasma convection pattern observed by the STARE radars during the passage of a westward traveling surge. The curved line in the figure represents the surge form as it passed over northern Scandinavia. One can see the counterstreaming plasma flow to the north and south of the auroral form. This pattern appears to be due to a region of negative space charge that is co-located with the surge form.

### DYNAMIC VARIATIONS OF CONVECTION

It seems well-established (Heppner, 1972a, 1977; Evans et al., 1980; Foster et al., 1981a, and others) that the effects of substorms are, for most time sectors: (a) an expansion of the region of sunward flow to lower latitudes and (b) a reversal in the convective flow. At times, the expansion extends below  $60^\circ\Lambda$  in the dusk sector. Some measurements have indicated that the amount of expansion is of the order of  $2^\circ$  for each unit increase in  $K_p$ . In the dawn sector, the expansion to lower latitudes is perhaps only about  $1^\circ$  for each unit increase in  $K_p$ .

The dynamic variability of the convection pattern is exemplified by the Chatanika radar observations presented in Figure 11-10. A single 24-hour interval of convection observations was reduced to the latitudinal variation of the electrostatic potential at half-hour intervals. When contours connecting latitudes at the same potential are drawn, the alternating expansion, contraction, and intensification of the pattern associated with temporal dynamics (substorms) during the observation interval are accentuated as large scale distortions of the contours.

The influence of substorms on the convective flow in the midnight sector seems less well-established at present. There is some evidence to suggest that the Harang discontinuity moves westwards at such times (Horwitz et al., 1978). However, Foster et al. (1981a, b) and Nielsen and Greenwald (1979) have observed that the discontinuity is nearly east-west aligned and that it moves equatorward in association with a substorm expansion. The equatorward motion causes an apparent westward shift of the convection reversal at a given latitude during a substorm expansion.

The dynamic variations in plasma convection near the dayside cusp are as yet not well understood.

### IONOSPHERIC CONVECTION TIMES

The observations discussed above clearly indicate that the high latitude ionosphere is in continual motion owing to convection electric fields. The effect that this motion has on the ionospheric densities and temperatures not only depends on the speed of the flow, but also on the path that a plasma flux tube follows. Figure 11-11 shows several plasma trajectories in a magnetic inertial frame for the case of a constant, cross-tail magnetospheric electric potential of 64 kV and a symmetric two-cell convection pattern. Note that this inertial reference frame is different from the corotational reference frame used for Figures 11-6 through 11-10. To transform the velocities observed in a corotational frame to the magnetic inertial frame it is necessary to add the observed convection velocity to the corotational velocity corresponding to the geographic latitude of the point under observation.

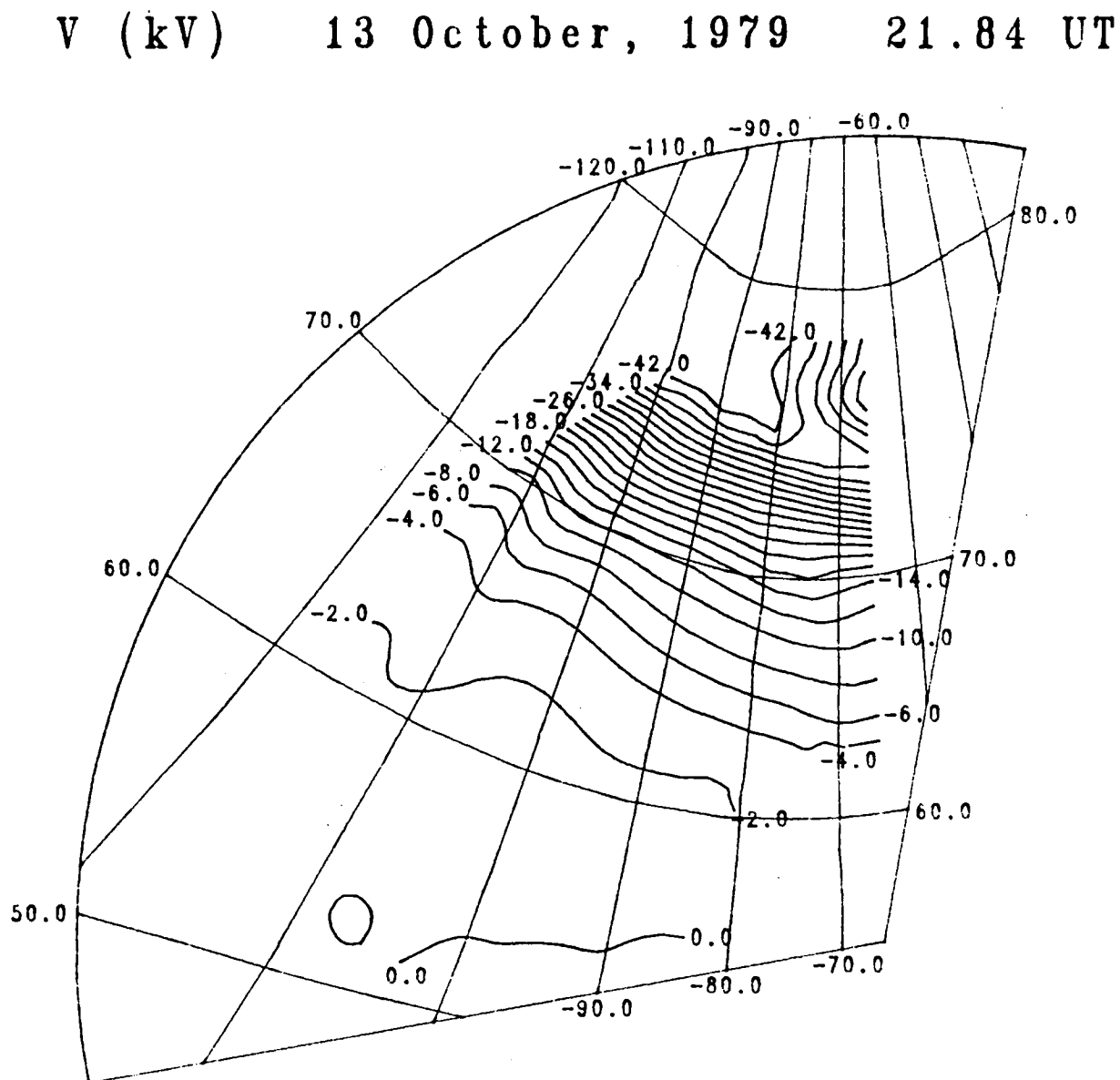


Figure 11-8. Maps of polarization electric field gathered during a substorm on 13 October 1979 by the scanning method. Plot shows polarization electric field contours as a function of apex latitude and longitude to the west and north of Millstone. Time at the center of the scan is 21.84 UT (16.29 EST).

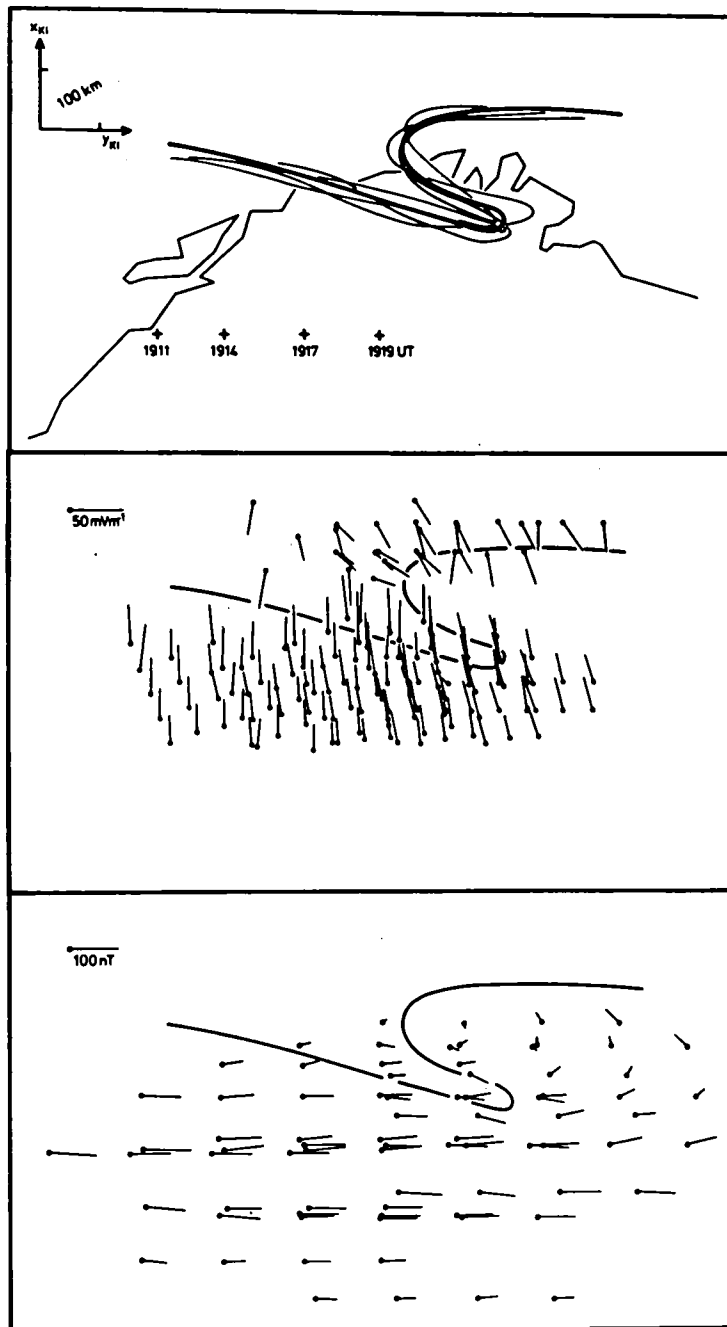


Figure 11-9. Electrodynamics of a westward traveling surge. The top panel illustrates the surge form at four epochs in its passage over northern Scandinavia. At each epoch the plus sign indicates the effective location of Kiruna at that time. The forms have been superimposed to illustrate the constancy of the shape, and all other measurements have been keyed to their location relative to the form. The middle panel illustrates the STARE electric field measurements, and the bottom panel shows the equivalent current measurements. From Inhester et al. (1981).

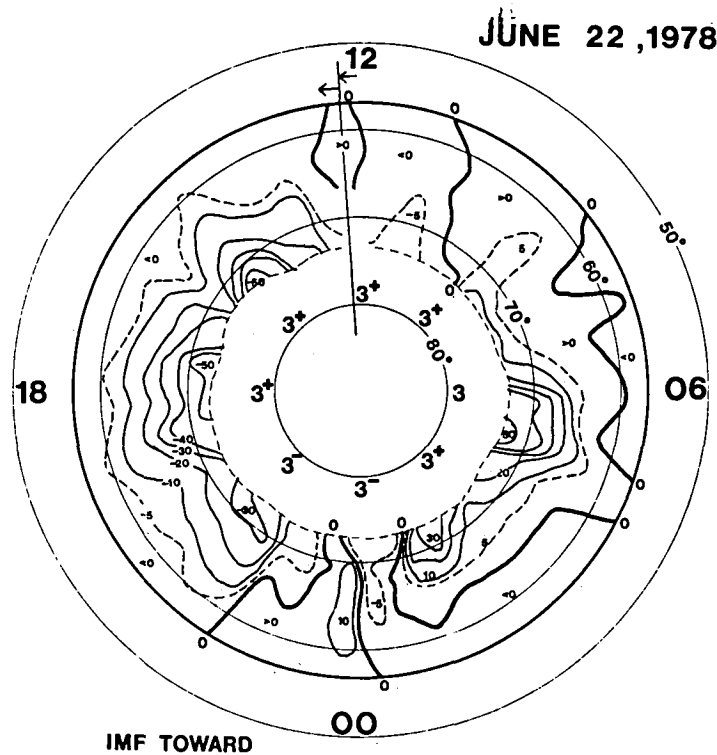


Figure 11-10. Contours of electric potential derived from radar ion drift data. Recurrent substorm activity produces the apparent latitudinal expansion and contraction of the potential field within the radar field of view. From Foster et al. (1982).

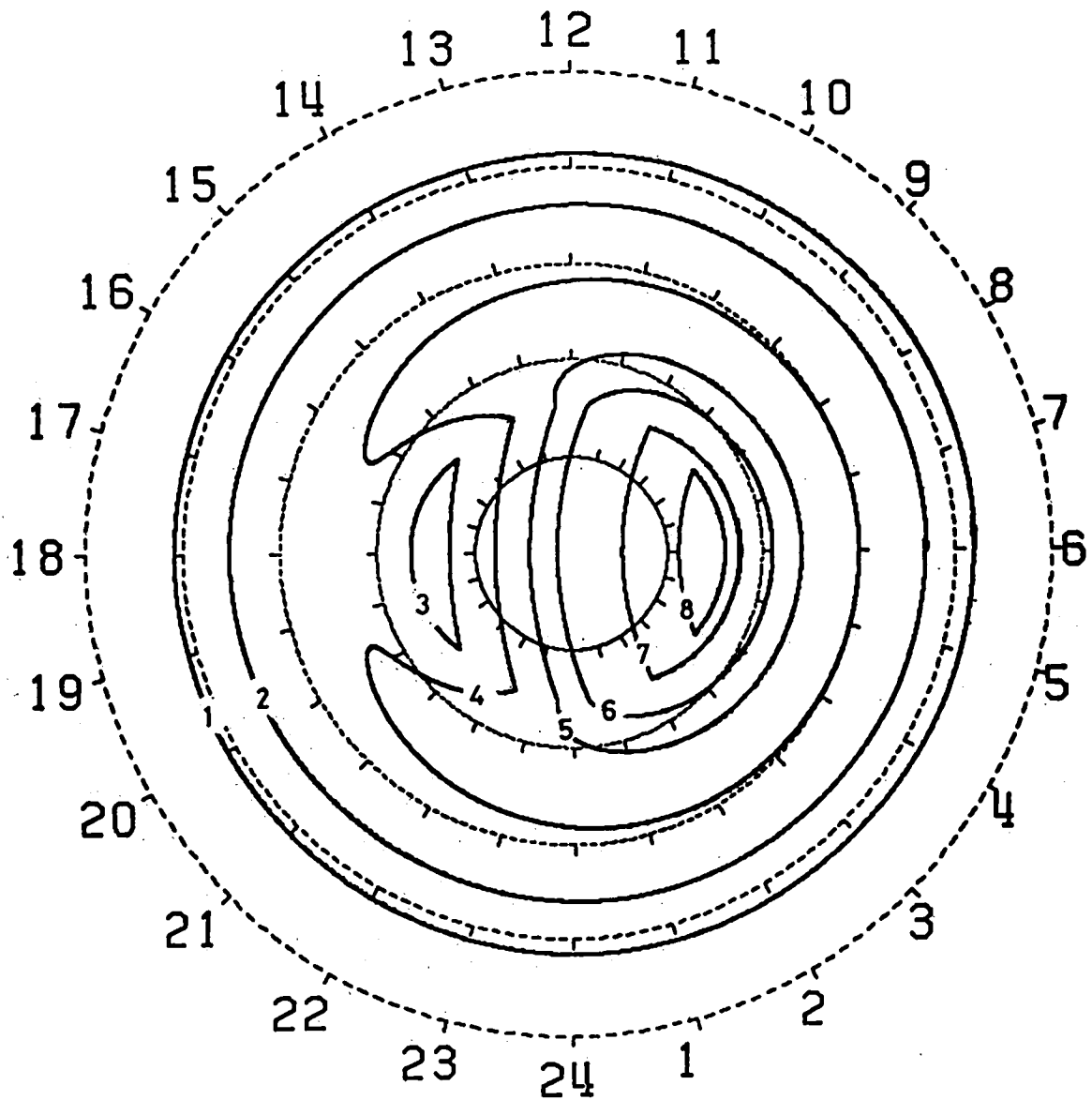
The time it takes for a flux tube of plasma to complete a trajectory depends not only on the length of the trajectory but also on the speed variation along the trajectory. Plasma flux tubes following the small circular trajectories in the center of the cells will circulate several times during the course of a day, while plasma flux tubes that pass near 1800 MLT take about 1.5 days to complete a trajectory. This latter result is due to the fact that in the dusk sector the westward (sunward) convective flow opposes the eastward corotational flow, with the result that a low speed region occurs near 1800 MLT equatorward of the auroral oval.

Several interesting ionospheric features can be anticipated from the plasma convection pattern. For example, if the low speed region near dusk is in darkness, the plasma will decay for a long time and low ionospheric densities will result. This is one of the principal mechanisms for the formation of the "main" electron density trough. Also, in the polar cap, the high-density dayside plasma convects across the terminator to the nightside. If the convection is slow, the plasma will decay significantly and a "polar hole" will form on the nightside just poleward of the auroral oval. On the other hand, if the plasma convection across the polar cap is rapid, an extended "tongue of ionization" will extend across the polar cap from the dayside to the nightside.

### IONOSPHERIC RESPONSE TO CONVECTION

During the last decade our understanding of the physical processes that control high-latitude ionospheric behavior has greatly increased. As a consequence, comprehensive large-scale models of the high-latitude ionosphere have been developed to describe the aurora, the E-region conductivity and currents, and the F region. The large-scale models of the high-latitude F region are capable of producing time-dependent, three-dimensional ion density distributions for six ion





LABEL	1	2	3	4	5	6	7	8
CIRCULATION PERIOD (day)	1.00	1.01	0.10	1.34	0.50	0.31	0.18	0.11

Figure 11-11. Plasma convection trajectories in a magnetic reference frame for a cross-tail magnetospheric electric potential of 64 kV. The trajectories have been numbered in order to indicate circulation times, which are tabulated in the lower part of the figure. From Sojka et al. (1981b).

species ( $\text{NO}^+$ ,  $\text{O}_2^+$ ,  $\text{N}_2^+$ ,  $\text{O}^+$ ,  $\text{N}^+$ ,  $\text{He}^+$ ) and electrons. The models take account of numerous high-latitude processes, including field-aligned diffusion, thermospheric winds, electrodynamic drifts, polar wind escape, energy dependent chemical reactions, magnetic storm induced neutral composition changes, thermal conduction, diffusion-thermal heat flow, local heat sources and sinks, and ion production due to solar EUV radiation and energetic particle precipitation (Roble and Rees, 1977; Knudsen et al., 1977; Watkins, 1978; Schunk and Raitt, 1980; Sojka et al., 1981a; Schunk and Sojka, 1982).

One of the early results to emerge from the use of these large-scale models is that high-latitude ionospheric features, such as the 'main trough', the 'ionization hole', the 'tongue of ionization', the 'aurorally produced ionization peaks', and the 'universal time effects', are a natural consequence of the competition between the various chemical and transport processes known to be operating in the high-latitude ionosphere. In addition, the model studies have shown that the electron density at F-region altitudes displays a significant variation with season, solar cycle, convection pattern, and universal time.

Several interesting ionospheric features occur in winter for weak convection (Figure 11-12). A deep ionization hole occurs in the dawn sector poleward of the auroral oval. The hole is a result of slow antisunward convection in combination with ordinary ionic recombination. At 300 km, the electron density in the hole can be as low as  $5 \times 10^2 \text{ cm}^{-3}$ . However, a short horizontal distance away, in the auroral oval, the densities can be more than three orders of magnitude larger,  $5 \times 10^5 \text{ cm}^{-3}$ . Another interesting ionospheric feature is the main or mid-latitude electron density trough. This trough, which is situated just equatorward of the nocturnal auroral oval, is a region of low electron density that has a narrow latitudinal extent but is extended in longitude. For winter solstice and slow convection, the electron density in the trough can be as low as  $10^3 \text{ cm}^{-3}$  at 300 km.

When the plasma convection pattern changes, the high-latitude F region can undergo major changes in a relatively short time. For example, for strong convection in winter (Figure 11-13), the increased antisunward plasma drift sweeps the daytime high density plasma into the polar cap, producing an extended tongue of ionization. This tongue of ionization fills in the polar hole, with the electron density at 300 km increasing from  $5 \times 10^2 \text{ cm}^{-3}$  to about  $10^5 \text{ cm}^{-3}$ . Likewise, for strong convection the density in the main trough is increased by more than an order of magnitude. Also, in summer the electron density in the high-latitude F region is typically more than an order of magnitude greater than in winter (Figure 11-14). This results because most of the polar cap is sunlit in summer, while most of it is in darkness in winter.

There are some interesting thermal features in the high-latitude F region. Elevated electron temperatures are found in the polar cusp and nocturnal auroral zone (4000-8000K). Also, ion temperature hot spots can appear in the dusk and/or dawn sectors depending on the plasma convection pattern (Figure 11-15). The hot spots are associated with strong plasma convection cells and the elevated ion temperatures are a result of an increased ion-neutral frictional heating rate. In the hot spot, the ion temperature is greatest at low altitudes ( $T_i > 4000\text{K}$ ) and decreases with altitude throughout the F region. This behavior is in sharp contrast with that found outside of the hot spot and at middle and low latitudes, where the ion temperature increases with altitude in the F region.

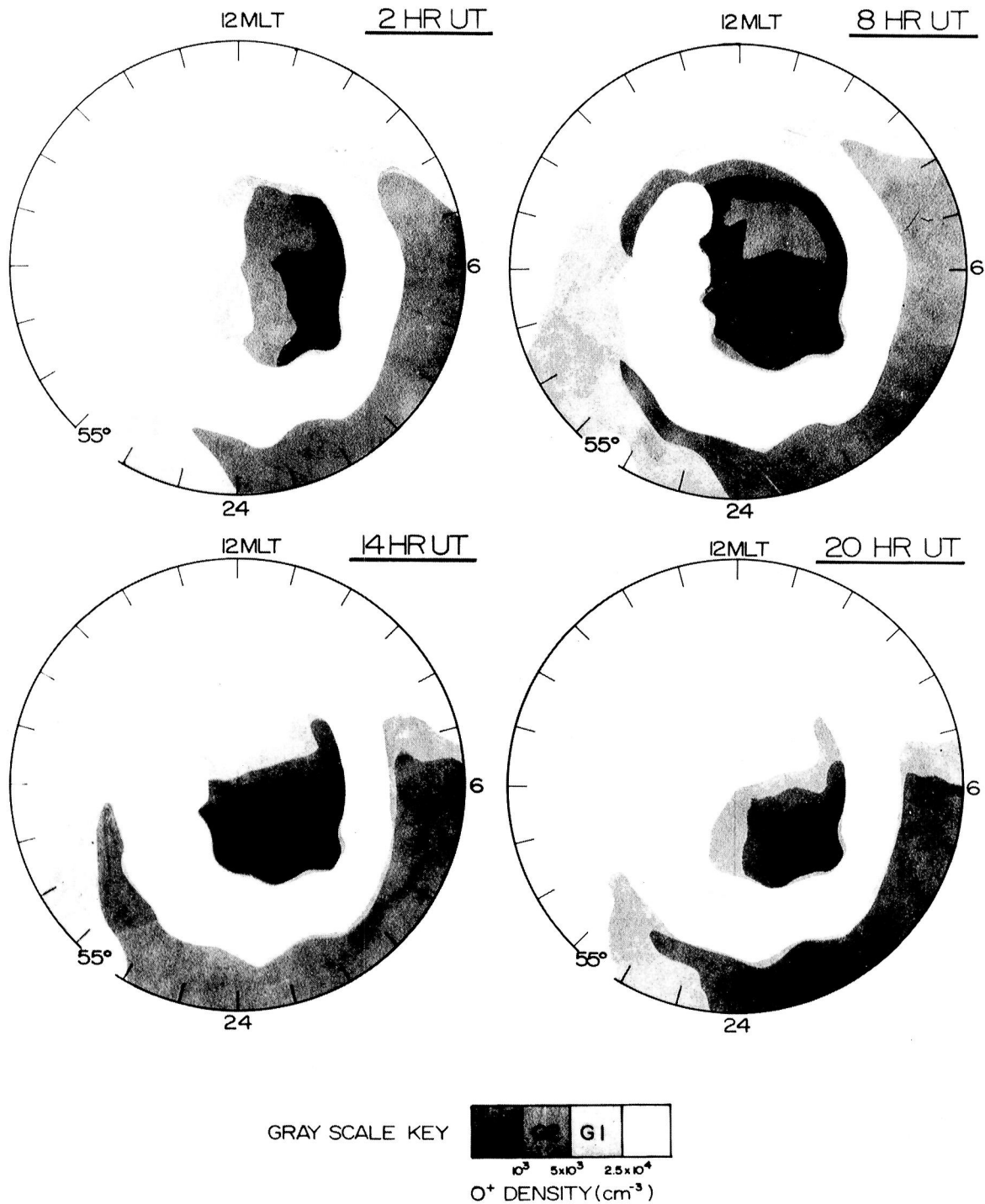


Figure 11-12.  $O^+$  density contours at 300 km for weak convection in winter. The contours are shown at four universal times in a magnetic quasi-inertial reference frame. The contour range was chosen to highlight the low density regions. From Sojka et al. (1981a)

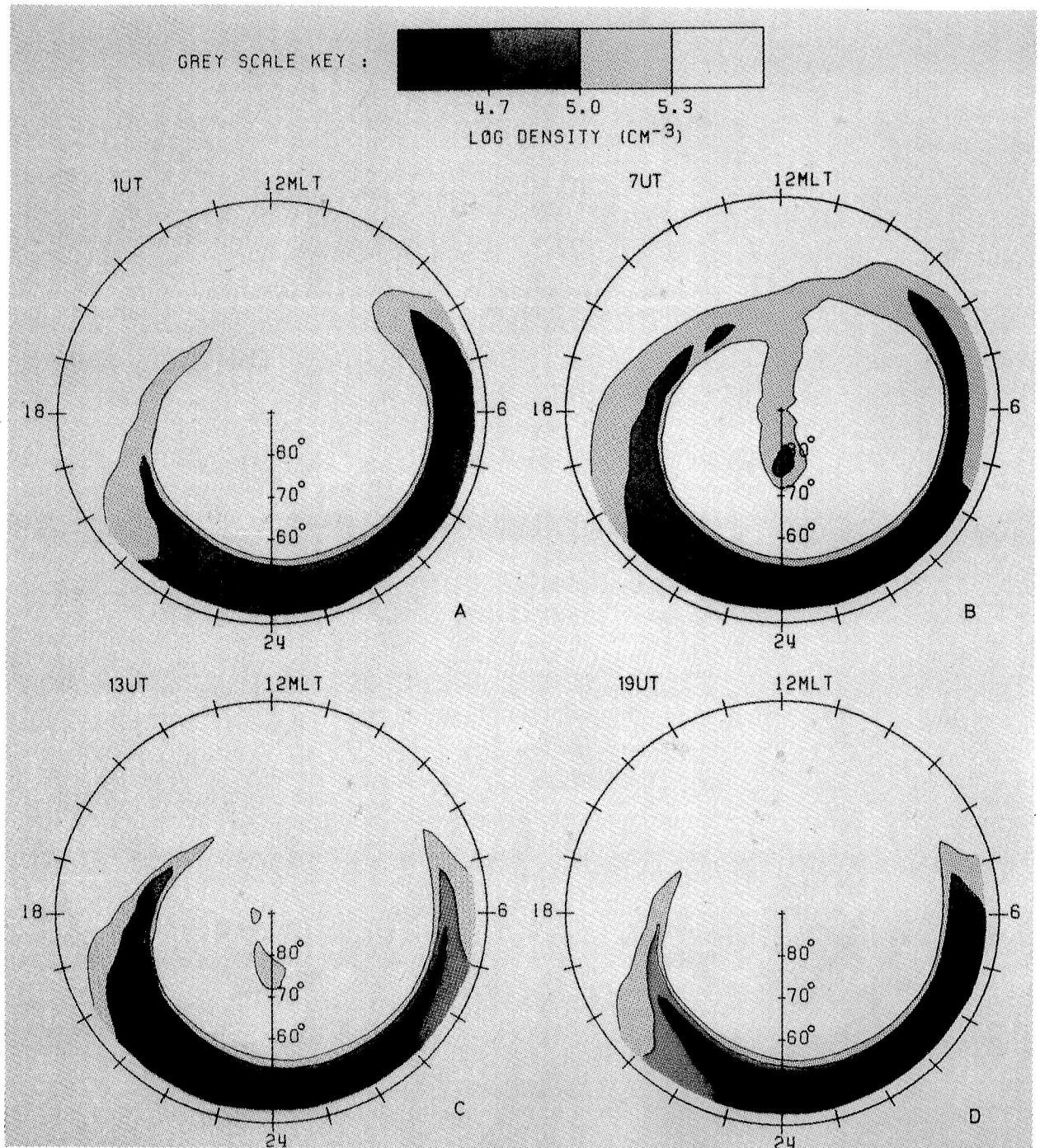


Figure 11-13.  $\text{O}^+$  density contours at 300 km for strong convection in winter. The contours are shown at four universal times in a magnetic quasi-inertial reference frame. From Sojka et al. (1981b).



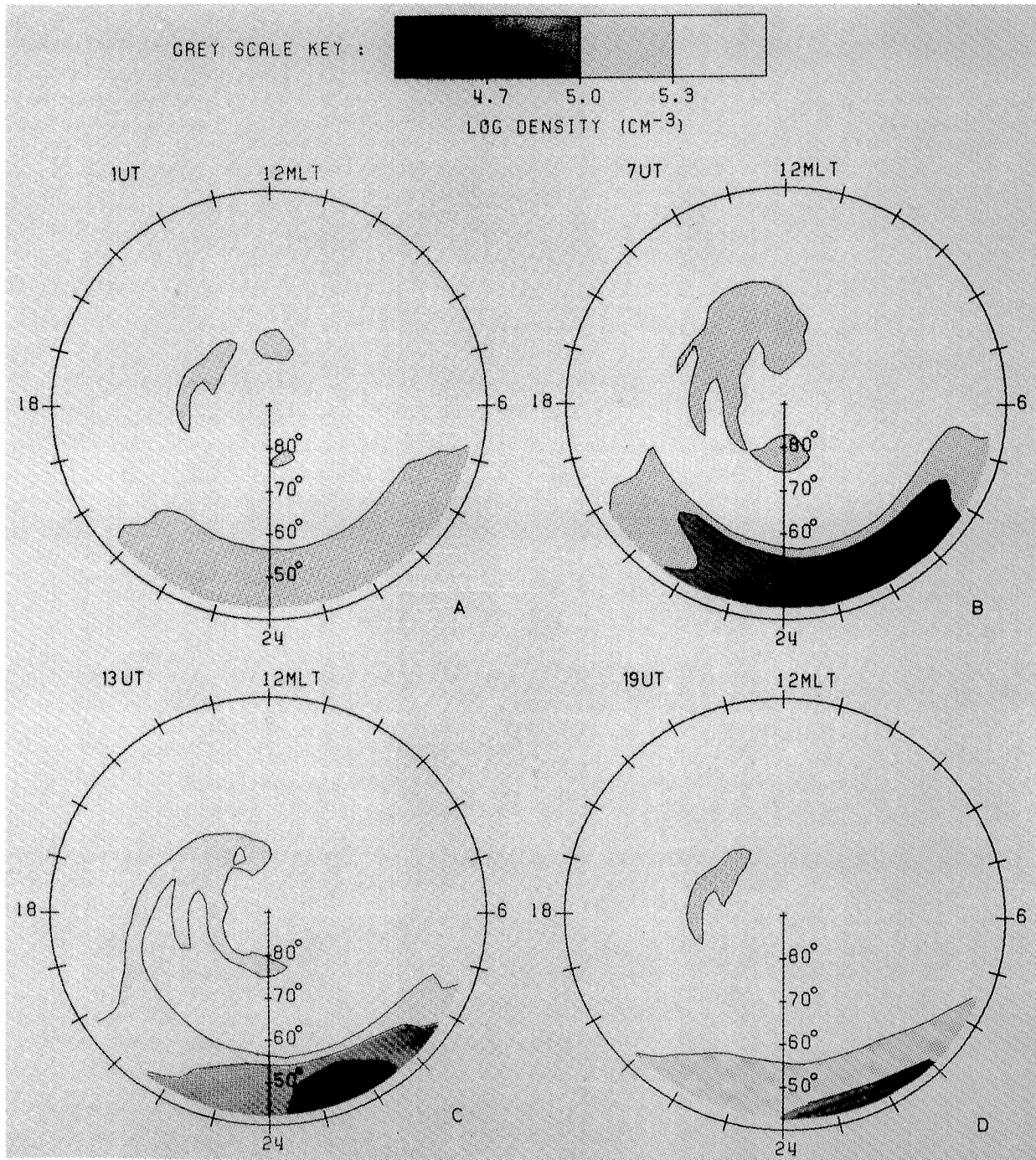


Figure 11-14.  $\text{O}^+$  density contours at 300 km for strong convection in summer. The contours are shown at four universal times in a magnetic quasi-inertial reference frame. From Sojka et al. (1982).

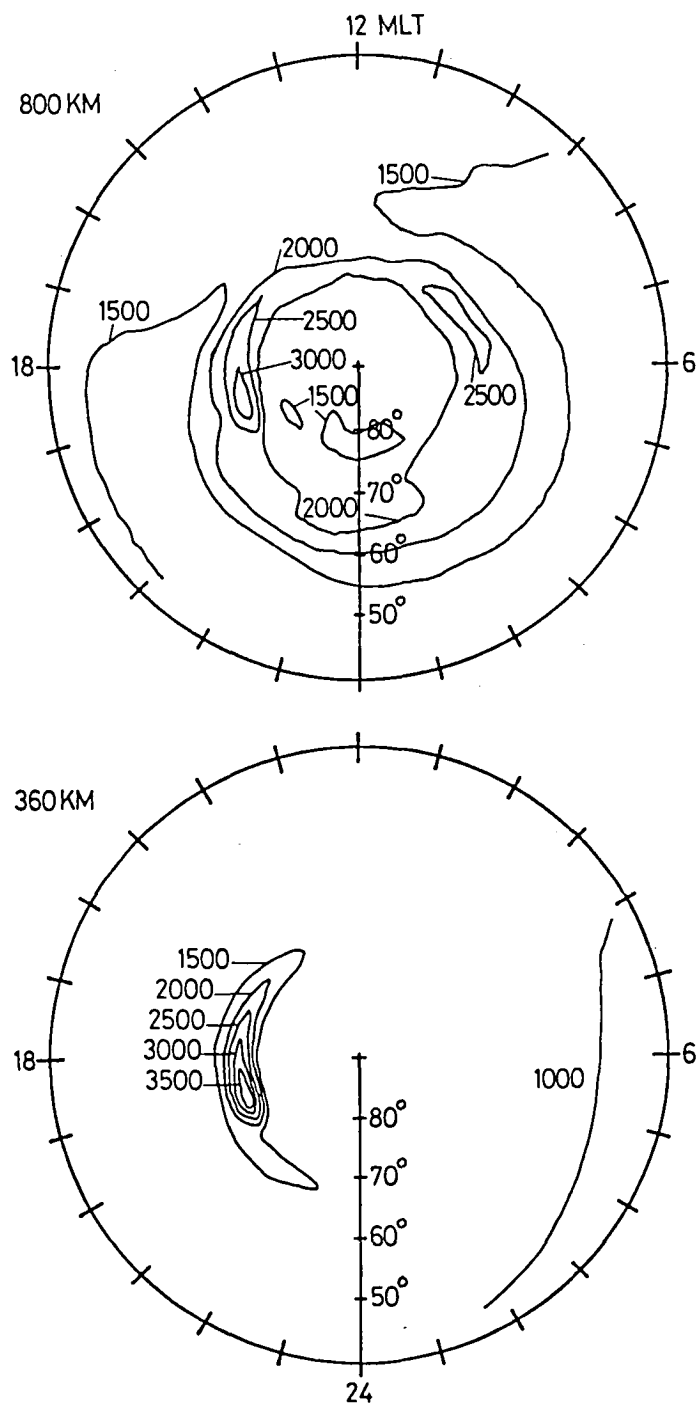


Figure 11-15. Contours of the ion temperature ( $^{\circ}\text{K}$ ) in a magnetic quasi-inertial reference frame for altitudes of 360 km (bottom panel) and 800 km (top panel). From Schunk and Sojka (1982).

Some theoretical work has been done on modeling magnetic storms (Sojka and Schunk, 1983). In particular, these authors studied the response of the high-latitude F region to an idealized magnetic storm for winter solstice and solar maximum conditions. The model storm was assumed to build in intensity over a 1-hour time period, persist at full intensity for about 1.5 hours, and then decay over a 3-hour time period. During the buildup, the auroral oval expands, plasma convection velocities increase, and particle precipitation is significantly enhanced. The main result obtained from this study was that the ionospheric response time is a strong function of altitude. The lower F region responds on a time scale of only minutes to the storm associated changes in the auroral precipitating electron flux, owing to the dominance of chemistry production-loss mechanisms over transport processes. At higher altitudes, in the vicinity of  $h_m F_2$ , the chemistry is balanced by both plasma diffusion along field lines and horizontal plasma convection, which acts to prolong the effect of the storm for many hours after it has ceased. The peak density responds only slowly to increased precipitation and may not reach its maximum enhanced value until over an hour after the storm main precipitation has passed. However, the F-region peak can be drastically altered on a time scale of minutes if large vertical transport velocities are associated with the storm electric field distribution. In the topside ionosphere the density variations are not correlated with the morphology of the storm auroral precipitation or the temporal variation of the storm electric field pattern. Time delays of up to 3 or 4 hours occur at high altitudes for "peak" densities to be reached after a storm, and the subsequent recovery is on the order of 5 hours. These long delays at altitudes above 400 km reflect the long time constants associated with plasma diffusion from low altitudes, where the plasma is created, to high altitudes.

## DIRECTION OF FUTURE RESEARCH

There is a need to establish the major physical processes that control convection, and there is also a need to determine what causes the spatial and temporal variations of the pattern. Some of the specific items to be addressed are:

- The polar cap potential distribution.
- The time-dependent response of the global ionization distribution to convection.
- The existence of multicell convection patterns.
- The coupling of convection patterns between conjugate ionospheres, and the possible effect of the different neutral wind systems.

To attack these problems requires simultaneous measurements of ion drifts, electron densities, particle temperatures, and neutral winds with a high spatial and temporal resolution. The availability of such measurements would allow quantitative comparisons between theory and measurement, which in turn would provide insight into the physical processes that control high-latitude ionosphere-thermosphere coupling.

## V. POLAR WIND

At latitudes greater than  $60^\circ$  invariant, light ions ( $H^+$  and  $He^+$ ) are capable of escaping from the topside ionosphere owing to the pressure difference between the ionosphere and the magnetosphere. Because this flow is analogous to the solar wind, Axford (1968) coined the term 'polar wind' to describe it. The polar wind corresponds to a continual outflow of thermal ionospheric

ions that occurs over the entire high-latitude region, and therefore, it is an important source of magnetospheric ions.

## THEORETICAL MODELS

Numerous theoretical models have been developed during the last decade to describe the polar wind, including hydrodynamic, hydromagnetic, kinetic, and generalized transport models (cf., Banks and Holzer, 1969; Raitt et al., 1975; Holzer et al., 1971; Lemaire and Scherer, 1973; Schunk and Watkins, 1982). The hydrodynamic, hydromagnetic and generalized transport equations are obtained by taking velocity moments of the Boltzmann equation in an effort to derive conservation equations for the physically significant moments of the distribution function, such as density, drift velocity, temperature, stress tensor and heat flow vector. The hydrodynamic equations result if the plasma is assumed to be collision-dominated, while the hydromagnetic equations correspond to the collisionless moment equations. The generalized transport equations are similar to the hydromagnetic equations, except that collisional terms are retained, and therefore these equations provide a continuous transition from the collision-dominated to the collisionless regimes. The kinetic models, on the other hand, are obtained by directly integrating the collisionless Boltzmann equation.

To date, the hydrodynamic model has received the most attention. With this model, altitude profiles of the ion and electron densities, drift velocities and temperatures are obtained over the altitude range of from 200-2000 km. This model has been used to study the effect that certain transport processes have on the polar wind, such as thermal diffusion, thermal conduction, diffusion-thermal heat flow, viscous stress, and convection electric fields. This model has also been used to study the seasonal, solar cycle, and geomagnetic activity variations of the polar wind. Based on these model studies, the polar wind is expected to have the following characteristics: (1) The flow could be either subsonic or supersonic; (2) The composition is dominated by the light ions  $H^+$  and  $He^+$ ; (3) the  $H^+$  escape flux should vary from  $10^7$  to  $5 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$  and the  $He^+$  escape flux from  $10^5$  to  $10^7 \text{ cm}^{-2} \text{ s}^{-1}$  at an altitude of 2000 km, depending on the geophysical conditions; (4) The ion temperatures should be less than 8000K (less than 1 eV); (5) There should be a significant seasonal variation of the  $He^+$  escape flux, with the escape flux in winter about 20-40 times larger than that in summer; and (6) There should be an appreciable solar cycle variation of the  $H^+$  escape flux.

At altitudes above 2000 km, the polar wind becomes collisionless, and consequently, the hydrodynamic model becomes invalid. At the altitudes where the polar wind is not collision-dominated, the ion pressure tensor is not isotropic and the ion heat flow is not described by a Spitzer conductivity. As noted earlier, the collisionless characteristics of the polar wind can be described by kinetic, hydromagnetic and generalized transport models. For supersonic flow, these models produce density and drift velocity profiles that are similar to those obtained from the hydrodynamic model. However, the temperature and heat flow profiles are significantly different. Specifically, these models predict appreciable ion and electron temperature anisotropies in the collisionless regime. For supersonic flow, the ion and electron temperatures parallel to the magnetic field,  $T_{\parallel}$ , are greater than the corresponding perpendicular temperatures,  $T_{\perp}$ . For  $H^+$ ,  $T_{\parallel}/T_{\perp} \approx 20$  at a distance of 10 Earth radii. For subsonic flow, on the other hand,  $T_{\parallel} < T_{\perp}$  for both ions and electrons if there is a heat flow from high to low altitudes. Also, in the collisionless regime, the ion and electron velocity distributions should be asymmetric, with elongated tails along the magnetic field in the upward direction for supersonic  $H^+$  outflow.

Figure 11-16 shows schematically how the  $H^+$  velocity distribution varies with altitude for a collisionless supersonic  $H^+$  outflow (cf., Holzer et al., 1971; Sojka et al., 1983). Starting with a drifting Maxwellian distribution at the exobase ( $\sim 4500$  km), the kinetic evolution of the  $H^+$  velocity distribution produces an anisotropic distribution at higher altitudes with an asymmetry



along the field-aligned direction due to the drift. These non-Maxwellian characteristics grow with altitude, and at altitudes greater than about  $10 R_E$ ,  $T_{\parallel}/T_{\perp} > 20$ . Also, above this altitude the asymmetry with respect to the field-aligned direction is sufficient to cause a displacement between the drift velocity and the peak of the velocity distribution. Such large deviations from a Maxwellian have important implications for the stability of the polar wind. In this regard, it should be noted that only a small field-aligned asymmetry is expected for subsonic outflow.

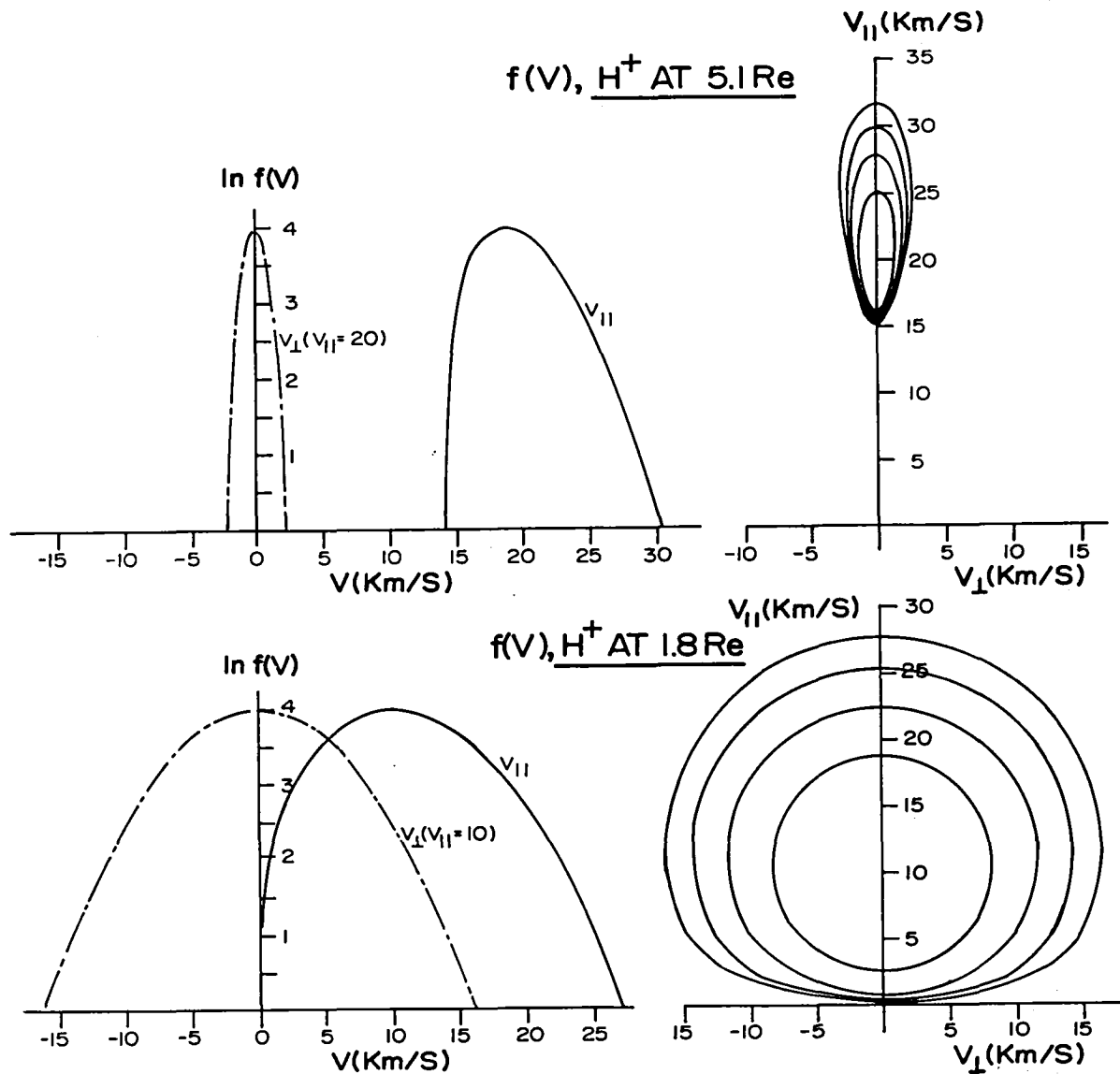


Figure 11-16. Schematic diagram showing the kinetic evolution of the  $H^+$  velocity distribution for a collisionless supersonic polar wind. Starting with a drifting Maxwellian at the exobase (lower panel), the distribution function at higher altitudes becomes asymmetric and anisotropic (upper panel) as the ions flow along the diverging magnetic field lines.

All of the polar wind characteristics described above were determined from steady-state models of ionospheric outflow. To date, very little effort has been devoted to studying the temporal behavior of the polar wind. Mitchell and Palmadesso (1983) used a generalized transport model to study the response of the polar wind to the sudden application of a large upward field-aligned current. These authors found that the plasma responds to the current on several time scales. After an initial rapid heating of the electrons due to precipitation in a converging magnetic field, the polar wind exhibits thermal oscillations which persist for about an hour.

Singh and Schunk (1982) have conducted numerical simulations that may be relevant to the initial expansion of the polar wind. Specifically, these authors studied the expansion of a collisionless, electrically neutral, multi-ion plasma into a vacuum. The study indicated that energetic ions are created when a plasma expands. For  $H^+$ , the typical energy gained during the expansion is about  $10 kT_e$ . Since  $T_e$  probably varies from about 1000 to 10,000 K in the high-latitude topside ionosphere, the energy of the  $H^+$  ions accelerated in an expanding polar wind should lie in the range of from 1 to 10 eV.

## OBSERVATIONS

Published data on the outflow of plasma from the Earth's ionosphere fall into two general categories: direct measurements of plasma flow parallel to the geomagnetic field; and measurements of ionospheric plasma parameters indirectly related to the effects of outflow. Direct measurements of the outflow velocity are particularly sparse, but even the published indirect measurements related to light ion outflow are limited in extent. Both the direct and indirect measurements are restricted to altitudes below about 3,000 km. At present, there are no magnetospheric measurements directly related to the polar wind, other than the existence of thermal and suprathermal plasma in the magnetospheric tail.

The existence of the polar wind was first determined directly by Hoffman (1970) using data from the Explorer-31 satellite. He studied the phase shift between  $O^+$  and  $H^+$  ions detected by an ion mass spectrometer mounted with a radial view direction on a satellite orbiting in a cartwheel mode (spin axis normal to the orbit plane). The orbital characteristics of Explorer-31 enabled measurements to be made above an altitude of about 500 km. The measurements showed the existence of  $H^+$  ion flow velocities of up to 10 km/s parallel to the geomagnetic field and flux levels of the order of  $10^8 \text{ cm}^{-2} \text{ sec}^{-1}$  above 2,500 km. A similar ion mass spectrometer to analyze ionospheric ions was flown on the ISIS-2 satellite in a 1,400 km circular, polar orbit. During the lifetime of the satellite, a large data base of observations over both the southern and northern polar regions was obtained (Hoffman et al., 1974; Maier and Hoffman, 1974; Hoffman and Dodson, 1980).

The direct observations of the polar wind have confirmed some of the predictions of the hydrodynamic models. The satellite measurements of the light ion ( $H^+$  and  $He^+$ ) densities, outflow velocities, and escape fluxes are in good qualitative agreement with the theoretical predictions. The seasonal and solar cycle variations of the light ion escape fluxes have also been confirmed. In addition, direct measurements have indicated that the  $H^+$  temperature in the polar wind is frequently greater than the  $O^+$  temperature, with the  $H^+/O^+$  temperature ratio generally in the range of 1-3.

Indirect evidence for the polar wind has been obtained from in situ ion density measurements made by satellites in high inclination, circular orbits. Typically, above 600 km, a satellite in such a circular orbit will observe a marked latitudinal variation of the  $H^+$  and  $He^+$  densities, with more than an order of magnitude decrease in the ion densities occurring as the satellite moves poleward near  $60^\circ$  (invariant latitude). The light ion densities are generally depressed over the entire polar region, which implies a spatially extensive light ion outflow.

## DIRECTION OF FUTURE RESEARCH

Although some of the predictions of polar wind models have been confirmed, many important characteristics still have to be verified. Owing to the lack of simultaneous measurements of ion drift velocities and temperatures, it is still not known whether the light ion outflow is subsonic or supersonic, i.e., whether it is a polar breeze or a polar wind. The predicted variations of the polar wind densities and temperatures with variations of the convection electric field and the  $H^+$  escape velocity have not been confirmed. In addition, there are no measured altitude profiles of the relevant polar wind parameters, such as density, drift velocity and temperature. Furthermore, there are no measurements bearing on the anisotropy or asymmetry of the ion distribution functions. Therefore, in the future the most important aspect of polar wind research will involve experimental testing of the various theoretical predictions. With regard to future modeling efforts, it is important to study the temporal evolution of the polar wind in much greater detail and to determine whether or not suprathermal ions can be created as the plasma expands in response to density and temperature enhancements. Also, it is important to study the stability of the polar wind.

## VI. CONCLUSIONS

To date, the different elements of the ionosphere-thermosphere system have generally been studied separately. That is, there have been many studies dealing with high-latitude currents and conductivities, thermospheric winds, plasma convection, and the polar wind, but the time-dependent coupling between the different elements has largely been ignored. Therefore, in the future, it is important to study the coupled system for a range of time scales, covering wave phenomena, diurnal, seasonal, and solar cycle variations.

In the course of studying the different elements separately, a number of major unresolved problems have surfaced. However, many of these problems may simply be related to the lack of consideration of the coupled, time-dependent behavior of the ionosphere-thermosphere system. At the present time, some of the major unresolved problems are the following:

- What is the relationship between the Birkeland currents, horizontal currents, and the convection electric field?
- What is the relationship between ionospheric convection and thermospheric winds? Does the wind drag or drive the ionosphere?
- Can the atmosphere-ionosphere dynamo affect magnetospheric transport?
- Can atmospheric gravity waves and magnetospheric MHD waves affect the ionospheric energy budget?
- How do the potential distributions and plasma convection patterns vary both spatially and temporally?
- What is the response of the global ionization distribution to plasma convection?

- What is the relationship of convection patterns in the conjugate ionospheres and neutral wind systems?
- Is the polar wind supersonic or subsonic?
- Is the polar wind stable?
- What fraction of magnetospheric ions is related to the polar wind?

## VII. REFERENCES

- Axford, W. I., The polar wind and the terrestrial helium budget, J. Geophys. Res., **73**, 6855, 1968.
- Banks, P. M., Ionosphere-magnetosphere coupling. 2. Electric fields, Rev. Geophys. Space Phys., **13**, 874, 1975.
- Banks, P. M., and T. E. Holzer, High-latitude plasma transport: The polar wind, J. Geophys. Res., **74**, 6317-6332, 1969.
- Banks, P. M., and J. R. Doupnik, A review of auroral zone electrodynamics deduced from incoherent scatter radar observations, J. Atmos. Terr. Phys., **37**, 951, 1975.
- Baumjohann, W., J. Untiedt, and R. A. Greenwald, Joint two-dimensional observations of ground magnetic and ionospheric electric fields associated with auroral zone currents. 1. Three-dimensional current flows associated with a substorm-intensified eastward electrojet, J. Geophys. Res., **85**, 1963, 1980.
- Doupnik, J. R., P. M. Banks, M. J. Baron, C. L. Rino, and J. Petriceks, Direct measurements of plasma drift velocities at high magnetic latitudes, J. Geophys. Res., **77**, 4268, 1972.
- Evans, J. V., J. M. Holt, W. L. Oliver, and R. H. Wand, Millstone Hill incoherent scatter observations of auroral convection over  $60^\circ < \Lambda < 75^\circ$ . 2. Initial results, J. Geophys. Res., **85**, 41, 1980.
- Foster, J. C., An empirical electric field model derived from Chatanika radar data, J. Geophys. Res., in press, 1983.
- Foster, J. C., J. R. Doupnik, and G. S. Stiles, Large-scale patterns of auroral ionospheric convection observed with the Chatanika Radar, J. Geophys. Res., **86**, 11357-11371, 1981a.
- Foster, J. C., J. R. Doupnik, and G. S. Stiles, Ionospheric convection and currents in the midnight sector on November 8, 1979, J. Geophys. Res., **86**, 2143-2148, 1981b.
- Foster, J. C., P. M. Banks, and J. R. Doupnik, Electrostatic potential in the auroral ionosphere derived from Chatanika radar observations, J. Geophys. Res., **87**, 7513-7524, 1982.
- Greenwald, R. A., Electric fields in the ionosphere and magnetosphere, Space Sci. Rev., **34**, 305-315, 1983.

- Greenwald, R. A., and A. D. M. Walker, Energetics of long period resonant hydromagnetic waves, Geophys. Res. Lett., **10**, 745-748, 1980.
- Greenwald, R. A., W. Weiss, E. Nielsen, and N. R. Thomson, Stare: A new radar auroral backscatter experiment in northern Scandinavia, Radio Sci., **13**, 1021, 1978.
- Haerendel, G. R., R. Lust, and E. Reiger, Motion of artificial ion clouds in the upper atmosphere, Planet. Space Sci., **15**, 1, 1967.
- Hays, P. B., R. A. Jones, and M. H. Rees, Auroral heating and the composition of the neutral atmosphere, Planet. Space Sci., **21**, 559-573, 1973.
- Heelis, R. A., and W. B. Hanson, High-latitude ion convection in the nighttime F region, J. Geophys. Res., **85**, 1995-2002, 1980.
- Heelis, R. A., W. B. Hanson, and J. L. Burch, Ion convection velocity reversals in the dayside cleft, J. Geophys. Res., **81**, 3803, 1976.
- Heelis, R. A., J. K. Lowell, and R. W. Spiro, A model of the high-latitude ionospheric convection pattern, J. Geophys. Res., **87**, 6339-6345, 1982.
- Heppner, J. P., Electric field variations during substorms, Planet. Space Sci., **20**, 1475, 1972a.
- Heppner, J. P., Polar cap electric field distributions related to the interplanetary magnetic field direction, J. Geophys. Res., **77**, 4877, 1972b.
- Heppner, J. P., Empirical models of high-latitude electric fields, J. Geophys. Res., **82**, 1115, 1977.
- Hoffman, J. H., Studies of the composition of the ionosphere with a magnetic deflection mass spectrometer, Int. J. Mass Spectrom. Ion Phys., **4**, 315-322, 1970.
- Hoffman, J. H., and W. H. Dodson, Light ion concentrations and fluxes in the polar regions during magnetically quiet times, J. Geophys. Res., **85**, 626, 1980.
- Hoffman, J. H., W. H. Dodson, C. R. Lippincott, and H. D. Hammack, Initial ion composition results from the ISIS-2 satellite, J. Geophys. Res., **79**, 4246, 1974.
- Holzer, T. E., J. A. Fedder, and P. M. Banks, A comparison of kinetic and hydrodynamic models of an expanding ion-exosphere, J. Geophys. Res., **76**, 2453, 1971.
- Horwitz, J. L., J. R. Doupnik, P. M. Banks, Y. Kamide, and S.-I. Akasofu, The latitudinal distributions of auroral zone electric fields and ground magnetic perturbations and their response to variations in the interplanetary magnetic field, J. Geophys. Res., **83**, 2071, 1978.
- Inhester, B., W. Baumjohann, R. A. Greenwald, and E. Nielsen, Joint two-dimensional observations of ground magnetic and ionospheric electric fields associated with auroral zone currents, 3, Auroral zone currents during the passage of a westward travelling surge, J. Geophys. Res., **49**, 155, 1981.

- Kamide, Y., and S. Matsushita, Simulation studies of ionospheric electric fields and currents in relation to field aligned currents. 1. Quiet periods, J. Geophys. Res., **84**, 4083, 1979a.
- Kamide, Y., and S. Matsushita, Simulation studies of ionospheric electric fields and currents in relation to field aligned currents. 2. Disturbed periods, J. Geophys. Res., **84**, 4099, 1979b.
- Kamide, Y., A. D. Richmond, and S. Matsushita, Estimation of ionospheric electric fields, ionospheric currents, and field-aligned currents from ground magnetic records, J. Geophys. Res., **86**, 801-813, 1981.
- Kelley, M. C., F. S. Mozer, and U. V. Fablesen, Electric fields in the nighttime and daytime auroral zone, J. Geophys. Res., **76**, 6054, 1971.
- Kelley, M. C., G. Haerendel, H. Kappler, F. S. Mozer, and U. V. Fablesen, Electric field measurements in a major magnetospheric substorm, J. Geophys. Res., **80**, 3181, 1975.
- Kisabeth, J. L., On calculating magnetic and vector potential fields due to large-scale currents in an infinitely conducting earth, in Quantitative Modeling of Magnetospheric Processes, edited by W. P. Olson, 473-498, AGU, Washington, D.C., 1979.
- Knudsen, W. C., P. M. Banks, J. D. Winningham, and D. M. Klumpar, Numerical model of the convecting F2 ionosphere at high latitudes, J. Geophys. Res., **82**, 4784, 1977.
- Lemaire, J., and M. Scherer, Kinetic models of the solar and polar winds, Rev. Geophys. Space Phys., **11**, 427-468, 1973.
- Maier, E. J., and J. H. Hoffman, Observation of a two-temperature ion energy distribution in regions of polar wind flow, J. Geophys. Res., **79**, 2444, 1974.
- Maynard, N. C., Electric field measurements across the Harang discontinuity, J. Geophys. Res., **79**, 4620, 1974.
- Mitchell, H. G., and P. J. Palmadesso, A dynamic model for the auroral field line plasma in the presence of field-aligned current, J. Geophys. Res., **88**, 2131-2139, 1983.
- Mozer, F. S., and P. Lucht, The average auroral zone electric field, J. Geophys. Res., **79**, 1001, 1974.
- Neilsen, E., and R. A. Greenwald, Electron flow and visual aurora at the Harang discontinuity, J. Geophys. Res., **84**, 4189-4200, 1979.
- Nisbet, J. S., M. J. Miller, and L. A. Carpenter, Currents and electric fields in the ionosphere due to field-aligned auroral currents, J. Geophys. Res., **83**, 2647, 1978.
- Oliver, W. L., J. M. Holt, R. H. Wand, and J. V. Evans, Millstone Hill incoherent scatter observations of auroral convection over  $60^\circ < \Lambda < 75^\circ$ , Average patterns versus Kp, J. Geophys. Res., **88**, 5505, 1983.
- Potemra, T. A., Current systems in the earth's magnetosphere, Review of United States progress for the 1975-1978 IUGG Quadrennial Report, Rev. Geophys. Space Phys., **17**, 640, 1979.

- Raitt, W. J., R. W. Schunk, and P. M. Banks, A comparison of the temperature and density structure in high and low speed thermal proton flows, Planet. Space Sci., **23**, 1103-1117, 1975.
- Rees, D., T. J. Fuller-Rowell, and R. W. Smith, Measurements of high latitude thermospheric winds by rocket and groundbased techniques and their interpretation using a three-dimensional, time-dependent dynamical model, Planet. Space Sci., **28**, 919, 1980.
- Ree, M. H., B. A. Emery, R. G. Roble, and K. Stamnes, Neutral and ion heating by auroral electron precipitation, J. Geophys. Res., **88**, in press, 1983.
- Rishbeth, H., Thermospheric winds and the F-region: A review, J. Atmos. Terr. Phys., **34**, 1, 1972.
- Rishbeth, H., and J. C. G. Walker, Directional currents in nocturnal E-region layers, Planet. Space Sci., **30**, 209, 1982.
- Roble, R. G., and M. H. Rees, Time-dependent studies of the aurora: Effects of particle precipitation on the dynamic morphology of ionospheric and atmospheric properties, Planet. Space Sci., **25**, 991, 1977.
- Roble, R. G., R. E. Dickinson, and E. C. Ridley, Global circulation and temperature structure of the thermosphere with high-latitude plasma convection, J. Geophys. Res., **87**, 1599, 1982.
- Schunk, R. W., and W. J. Raitt, Atomic nitrogen and oxygen ions in the daytime high latitude F region, J. Geophys. Res., **85**, 1255, 1980.
- Schunk, R. W., and J. J. Sojka, Ionospheric hot spot at high latitudes, Geophys. Res. Lett., **9**, 1045-1048, 1982.
- Schunk, R. W., and D. S. Watkins, Proton temperature anisotropy in the polar wind, J. Geophys. Res., **87**, 171-180, 1982.
- Singh, N., and R. W. Schunk, Numerical calculations relevant to the initial expansion of the polar wind, J. Geophys. Res., **87**, 9154-9170, 1982.
- Smith, M. F., D. Rees, and T. J. Fuller-Rowell, The consequences of high latitude particle precipitation on global thermospheric dynamics, Planet. Space Sci., **30**, 1259, 1982.
- Sojka, J. J., and R. W. Schunk, A theoretical study of the high latitude F region's response to magnetospheric storm inputs, J. Geophys. Res., **88**, 2112-2122, 1983.
- Sojka, J. J., W. J. Raitt, and R. W. Schunk, A theoretical study of the high-latitude winter F region at solar minimum for low magnetic activity, J. Geophys. Res., **86**, 609-621, 1981a.
- Sojka, J. J., W. J. Raitt, and R. W. Schunk, Plasma density features associated with strong convection in the winter high-latitude F region, J. Geophys. Res., **86**, 6908-6916, 1981b.
- Sojka, J. J., R. W. Schunk, and W. J. Raitt, Seasonal variations of the high-latitude F region for strong convection, J. Geophys. Res., **87**, 187-198, 1982.

- Sojka, J. J., R. W. Schunk, J. E. Johnson, H. Waite, and C. R. Chappell, Characteristics of thermal and suprathermal ions associated with the dayside plasma trough as measured by the Dynamics Explorer retarding ion mass spectrometer, J. Geophys. Res., in press, 1983.
- Stern, D. P., Large-scale electric fields in the earth's magnetosphere, Rev. Geophys. Space Phys., **15**, 156, 1977.
- Watkins, B. J., A numerical computer investigation of the polar F-region ionosphere, Planet. Space Sci., **26**, 559, 1978.



# CHAPTER 12

## HIGH LATITUDE IONOSPHERIC STRUCTURE

### WORKING GROUP MEMBERS

Sidney L. Ossakow, Chairman  
*Naval Research Laboratory*

William Burke  
*Air Force Geophysics Laboratory*

Herbert C. Carlson, Ex Officio  
*Air Force Geophysics Laboratory*

Peter Gary  
*Los Alamos National Laboratory*

Rod Heelis  
*University of Texas*

Michael Keskinen  
*Naval Research Laboratory*

Nelson Maynard  
*NASA/Goddard Space Flight Center*

Ching Meng  
*Johns Hopkins University*

Edward Szuszczewicz, Co-Chairman  
*Naval Research Laboratory*

James Vickrey  
*SRI International*

## CHAPTER 12

### HIGH LATITUDE IONOSPHERIC STRUCTURE

I.	Introduction	12-3
II.	Sources and Observations of High-Latitude Structure	12-7
III.	Plasma Instability Theory	12-18
IV.	An Emerging Picture	12-29
V.	Future Studies	12-31
VI.	References	12-33

## I. INTRODUCTION

### IONOSPHERIC STRUCTURE IN GENERAL

The Earth's ionosphere is an important element in solar-terrestrial energy transfer processes. As a major terrestrial sink for many solar and magnetospheric events, the ionosphere has characteristic features that can be traced to such seemingly remote phenomena as solar flares, radiation belt wave-particle interactions and magnetospheric substorms.

In considering the multitude of solar-terrestrial plasma interactions, it is important to recognize that the high-latitude ionosphere is not altogether a simple receptor of various energy deposition processes. The high-latitude ionosphere can play an active feedback role by controlling the conductivity at the base of far-reaching magnetic field lines and by providing a plasma source for the magnetosphere. Indeed, the role of the ionosphere during magnetospheric substorms is emerging as a topic for meaningful study in the overall picture of magnetospheric-ionospheric coupling.

The accessibility of the ionosphere provides the opportunity for rather detailed investigations with combinations of in situ diagnostics and remote sensing techniques. Thus it provides the plasma physics community exciting challenges in studying many of the dynamic and unstable ionospheric plasma states. Indeed, there are regions of the Earth's ionosphere that are more irregular than not, more dynamic than quiescent, and more unstable than previously appreciated.

The study of ionospheric structure, while traditionally focused on density irregularities, more generally includes irregularities in density, temperature, ion composition and fields. The study of irregularities in the ionospheric F region has divided itself naturally into three regimes: the low, middle and high geomagnetic latitudes. These regimes have fundamentally different sources for the irregularities, associated with the degree of coupling to higher altitude magnetospheric phenomena. The coupling is related, of course, to the orientation of the geomagnetic field as a function of magnetic latitude. At low latitudes, where the geomagnetic field tends to be horizontal, coupling to higher altitude magnetospheric disturbances is inhibited. At high latitudes the more vertical magnetic field promotes strong magnetospheric-ionospheric coupling.

The term "ionospheric irregularities" has become synonymous with the terms "fluctuations", "structure", and "striations". Under varying conditions, ionospheric irregularities can populate a broad range of latitudes, longitudes and altitudes. The irregularities are generally magnetic field aligned, i.e., there is little variation along the geomagnetic field. Thus, the irregularity wavelength perpendicular to the geomagnetic field is much smaller than its parallel wavelength and this two dimensional nature of the irregularities simplifies most theoretical analyses. There can be a diversity of irregularity scale sizes, spanning up to six orders of magnitude (tens of centimeters to hundreds of kilometers). The existence of such a broad spectrum of irregularities is further complicated by varying geophysical conditions and wavelength-dependent mechanisms. While our understanding at high latitudes is still primitive, the past few years have seen exciting advances in our understanding of low-latitude ionospheric irregularities (Ossakow, 1979; Fejer and Kelley, 1980; Ossakow et al., 1982) in the areas of equatorial spread-F and the equatorial electrojet. This advance has resulted from combined improvements in experimental, theoretical, and computer simulation techniques. The rapid development and meshing of these research tools in the past few years has made this a propitious time for unraveling some of the basic causal mechanisms of high latitude ionospheric irregularities.

## EQUATORIAL SPREAD-F IRREGULARITIES – A SUCCESS STORY AND A GUIDE FOR HIGH LATITUDES

Traditionally the term spread F comes from the range and frequency spread that appears on ionosonde height versus frequency diagrams of a highly structured F-region ionosphere. Now the term equatorial spread F (ESF) identifies the nighttime equatorial F-region condition in which irregularities span scale sizes ranging from hundreds of kilometers to fractions of a meter (as detected by radar, satellite, rocket, and scintillation measurements). Although ESF was discovered over four decades ago from ionosonde data (Booker and Wells, 1938), it is only recently that these phenomena have become well understood (see Ossakow, 1979; Fejer and Kelley, 1980; Ossakow, 1981; Kelley and McClure, 1981; Ossakow et al., 1982; Zalesak et al., 1982; Kelley et al., 1982a, b; Singh and Szuszczewicz, 1983).

Before we proceed to fill in on our ESF understanding, let us examine the basic ESF geometry as depicted in Figure 12-1.  $N(y)$  represents the background F-region zero order electron density as a function of altitude ( $y$ ). Gravity  $\underline{g}$  points down, the ambient geomagnetic field  $\underline{B}$  is horizontal and pointing north ( $z$ ), and  $\underline{k}$  represents a horizontal perturbation vector pointing westward ( $x$ ). The maximum in electron density defines the F-peak. The basic physical picture of ESF that emerges is as follows. After sunset, E-region molecular ions recombine and there is effectively no E-region conductivity to short out F-region irregularities. Due to recombination and electrodynamics (upward motion of the F-region plasma due to  $\underline{E} \times \underline{B}$ ) the bottomside F-region background electron density gradient begins to steepen. The steepening is due to plasma  $\underline{E} \times \underline{B}$  uplift caused by an eastward electric field (a downward motion of the neutral atmosphere will have the same effect). When the altitude of the F region is high enough and/or the bottomside

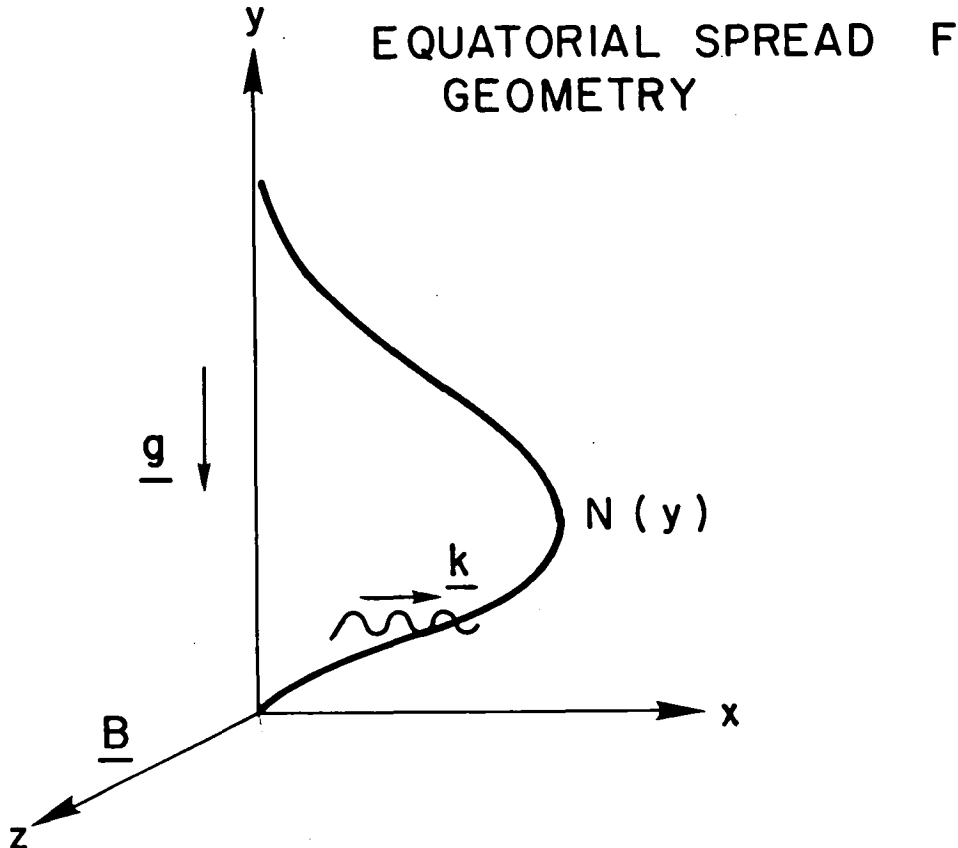


Figure 12-1. Basic equatorial spread F geometry.

background electron density gradient is steep enough to overcome recombination effects, long scale size plasma density fluctuations grow on the bottomside via the collisional Rayleigh-Taylor (R-T) instability mechanism (or possibly initiated by the  $E \times B$  gradient drift instability discussed in Section III). These irregularities will in turn form plasma density depletions (bubbles) on the bottomside which rise due to polarization induced  $\delta E \times B$  motion. The higher the F peak and/or the sharper the bottomside gradient, and the higher the percentage depletion, the faster the rise rate of the bubble and the faster the evolution of ESF. The topside of these steepening bubbles can bifurcate and form smaller scale structure by a cascade or multistep mechanism. An eastward neutral wind can cause the bubbles to move westward with respect to the bulk plasma motion. The long wavelength irregularities follow an approximate  $k^{-2+?}$  density fluctuation power law (Kelley et al., 1982a, b; Singh and Szuszczewicz, 1983). The above discussed picture of ESF development has been seen in the experimental observations and the theoretical and numerical simulation work and represents a rather intensive investigation of the phenomena. Thus it can be stated that major successes have been attained in the understanding of equatorial spread F.

The study of equatorial structure showed us the necessity for examining the entire wavelength spectrum of irregularities (from experimental, theoretical, and numerical simulation points of view) in order to unravel the causes of this structure. Similarly, other aspects of ESF phenomena will help us to increase our knowledge of high latitude irregularity phenomena. The tools developed at the equator are currently being applied at high-latitudes with specific similarities and differences detailed in subsequent sections.

## FOCUS ON HIGH-LATITUDE STRUCTURE

From a variety of experimental observations it is known that the high latitude ionosphere, from the auroral zone into the polar cap, is a nonequilibrium medium containing structure with scale sizes ranging from hundreds of kilometers to centimeters. The high-latitude ionosphere is profoundly affected by energy sources of magnetospheric origin. In the high-latitude ionosphere, several irregularity source mechanisms have been proposed: particle precipitation, bulk plasma processes, instability mechanisms, and neutral fluid dynamics. The relative importance of each mechanism undoubtedly depends on scale-size and geographic location with respect to the impressed magnetospheric boundaries.

It is logical to consider plasma processes in discussing high-latitude ionospheric irregularities since several sources of free energy are available to drive various plasma instabilities. Examples of these sources include density gradients, electric fields, neutral winds, velocity shears, and currents both parallel and perpendicular to the geomagnetic field. Both plasma macroinstabilities, which are fluid-like and operate on scales sizes  $\lambda \gg a_i$ , with  $a_i$  being the ion gyroradius, and plasma microinstabilities ( $\lambda \lesssim a_i$ ) have been invoked to account for high-latitude ionospheric structure. While the wide range of scale sizes makes for a richness of study, most theoretical analyses and numerical simulations cannot possibly handle this broad scale size spectrum all at once. As in the equatorial case the numerical problem has to be examined in pieces. Moreover, the very smallest scale sizes, gyroradius and below, must be treated via kinetic theory. These effects can be examined separately and then put back into the plasma fluid equations as anomalous transport coefficients. Thus, due to the diversity of scale sizes we use a two-step process for theoretical investigation. The inadequacy of linear instability theory to describe most aspects of ionospheric irregularities and their attendant effects leads us to study, by experimental, theoretical and numerical simulation techniques, the nonlinear evolution of the various irregularities. Because nonlinear plasma processes can couple structures in one scale-length regime to other scale-length regimes, the entire spectrum of irregularities must be studied.

Figure 12-2 depicts a model of the plasma enhancement used in the numerical simulations of Keskinen and Ossakow (1982) and are representative of one-type of high-latitude observation (Vickrey et al., 1980). This geometry is similar to that used in the ESF study (Figure 12-1), except that the magnetic field  $\underline{B}$  is now vertical instead of horizontal, and there is the field-aligned current,  $\underline{J}_{\parallel}$ . Such geometries with differences in drivers, are prime candidates for long-wavelength convective plasma fluid instabilities.

In studying high-latitude irregularities it is possible to witness, in the readily-accessible near-space environment, the evolution of structure in a magnetized, partially-ionized plasma that is subject to strong driving forces. And although the high-latitude situation is substantially more complicated than its equatorial counterpart, there are several straightforward directions to proceed in order to make substantive progress.

Important questions include:

1. When and where in the high-latitude ionosphere is plasma structure produced?
2. What are the scale size spectral distributions and associated causal mechanisms?
3. What is the role of plasma transport and what is the structure lifetime?
4. What are the roles of magnetospheric dynamics and the background ionosphere?, and finally;
5. How do plasma processes couple between altitude regimes?

Answers for some of these issues will be aided by studies of other working groups.

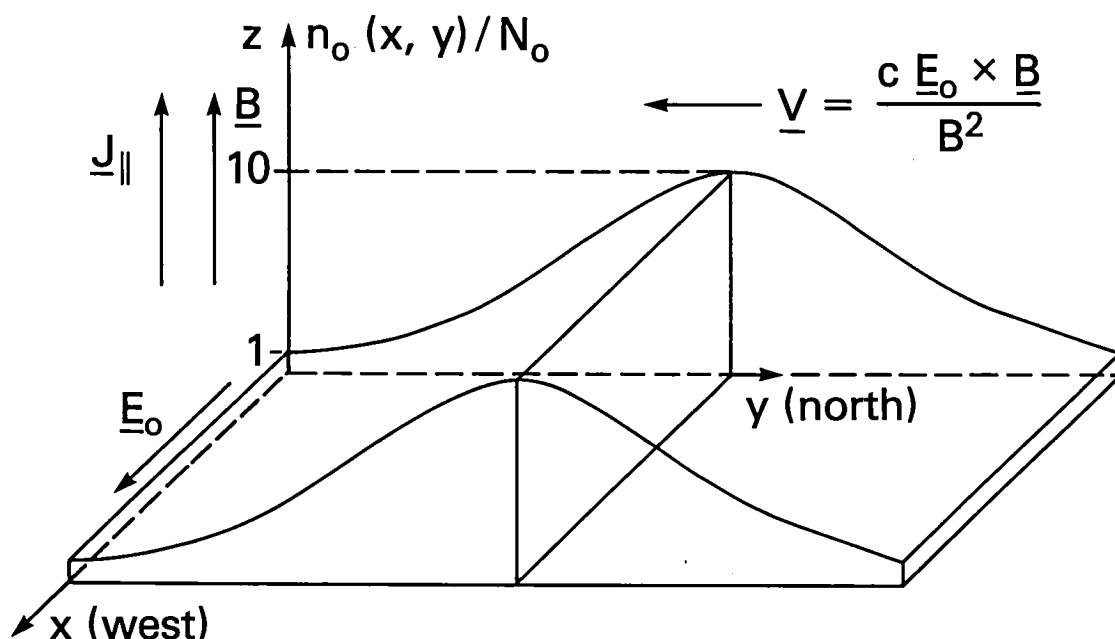


Figure 12-2. Model of plasma enhancement used in numerical simulations (taken from Keskinen and Ossakow, 1982). The quantity  $n_o(x,y)$  represents the initial plasma enhancement density profile while  $N_o$  is the average background auroral F-region plasma density.

To provide a guideline to a current understanding and future directions we have divided this report into five related sections. In Section II magnetospheric inputs and examples of ionospheric structure are presented. Section III outlines the theoretical plasma processes that are currently thought to be important in the high latitude ionosphere. Section IV assesses the present state of agreement between experiment and theory and presents an emerging picture of high-latitude plasma structure. Finally, Section V identifies directions for future studies.

## II. SOURCES AND OBSERVATIONS OF HIGH-LATITUDE STRUCTURE

Plasma in the high-latitude ionosphere is created primarily by solar EUV ionization and by precipitating electrons. The flux of precipitating ions is significantly less than that of the electrons and their contributions to ionospheric plasma will not be considered here. The plasma produced by solar EUV varies smoothly in density while that from precipitating electrons may be highly structured. There are many other possible sources of structure, reflecting a variety of plasma processes including instabilities driven by free energy sources such as field-aligned currents and magnetospheric electric fields.

Once created, these structures may be transported by impressed magnetospheric electric fields. They undergo further evolution through bifurcation and cascading processes that lead to ultimate dissipation. The purpose of this section is to explore the patterns of electron precipitation, convective electric fields, field-aligned currents, and the ionospheric density structure in the context of sources of irregularities and the structures themselves.

### ELECTRON PRECIPITATION STRUCTURES

It is generally accepted that precipitating electrons are a principal source of ionization structuring at high latitudes. In this, they play a dual role. First, they are free energy sources for certain plasma instabilities. Second, because magnetospheric electrons are inherently structured so too is the ionization they create.

Electron precipitation at high latitudes is variable in its energy characteristics and in its geographic distribution. This is illustrated in Figure 12-3. High-latitude precipitation is traditionally broken into three distinct regions, the auroral oval, the dayside cusp and the polar cap.

There are two characteristic types of electron precipitation in the auroral oval. Electrons from the central plasma sheet precipitate into the region marked diffuse aurora in Figure 12-3. These electrons typically have energies of a few keV and are isotropically distributed in pitch angle over the downcoming hemisphere. Although structured precipitation, e.g., that associated with inverted-V (Frank and Ackerson, 1971) and discrete auroral arcs, may be found in this region, electron precipitation is usually smoothly varying and produces diffuse auroral emission patterns.

Precipitation near the high latitude boundary of the auroral oval, sometimes referred to as boundary plasma sheet precipitation, is characterized by discrete structures. In quiet times the mean thermal energy of boundary plasma sheet electrons is typically a few hundred eV. In discrete arc- and inverted-V events these electrons are accelerated through field-aligned potential drops of several keV. During substorm periods the average energy of boundary plasma sheet electrons increases dramatically.

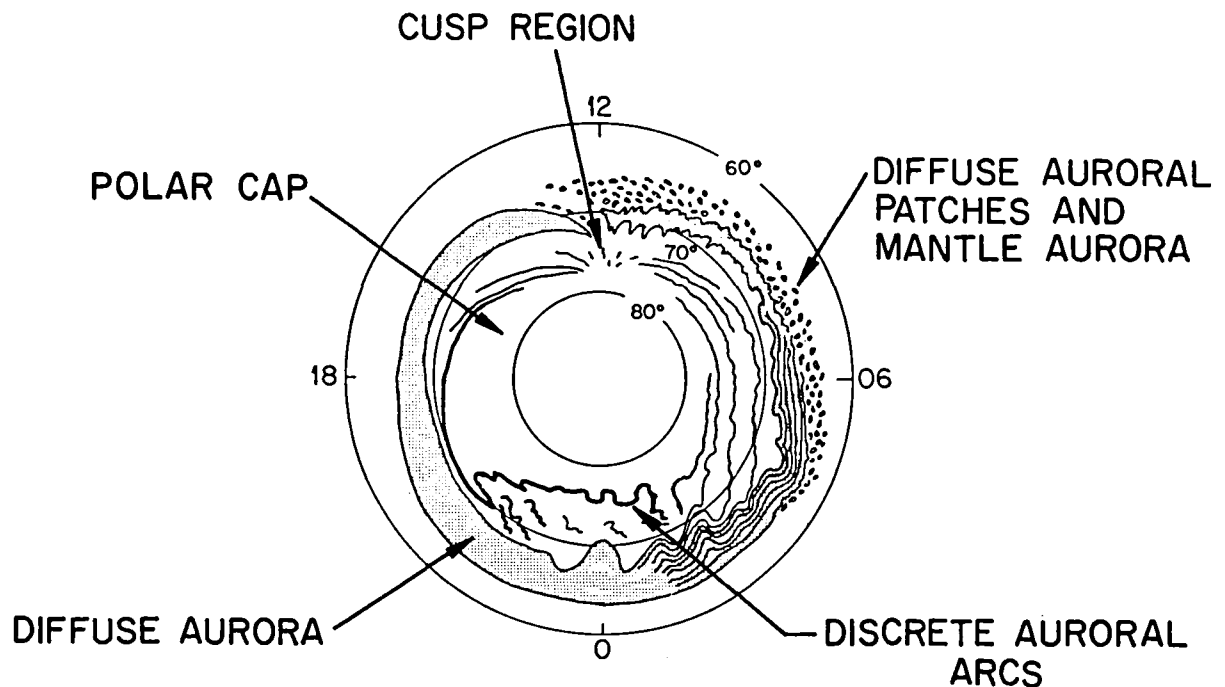


Figure 12-3. Phenomenological source term map of high latitude ionospheric structure (adapted from S. -I. Akasofu, 1981).

Returning to Figure 12-3, attention is directed to the most-poleward auroral form extending across the midnight sector. At approximately 2200 LT this arc has a distinctive kink, then extends smoothly into the evening/afternoon sector. Because of its apparent motion when viewed from the ground, the kink (characteristic of substorm periods) is known as a westward traveling surge. Precipitation near westward traveling surges is intense and spatially diverse. Most of the electrons have energies above 1 keV and create a large fraction of their ionization in the E region.

Near the equatorward boundary of the diffuse aurora, precipitating electrons have been detected (Tanskanen et al., 1981) with energies  $\sim 100$  eV at flux levels of  $10^{10}$   $\text{e}/(\text{cm}^2\text{-s-sr-keV})$ . Ionization from the electrons is produced at altitudes greater than 200 km. Recall that in the nightside auroral oval the precipitating particles are the primary sources of ionization and, hence, control the ionospheric conductivity.



The complexity of high-latitude electron precipitation is illustrated in Figure 12-4. Here we have plotted count rates measured in situ by an uplooking satellite-borne spectrometer in three energy channels near 0.27, 2.3, and 20 keV. The magnetic-latitude/magnetic-local-time trajectory of the satellite across the southern high-latitude ionosphere is provided in the figure for reference. The diffuse auroral precipitation on both the evening and morning sides is marked by substantial counts of electrons with energies of both 0.27 and 2.3 keV. As the satellite entered the boundary plasma sheet on the evening side the counts of 2.3 keV electrons fell to background levels. The counts in the 0.27 keV channel remained high and extremely varied up to magnetic latitudes of 84°. At magnetic latitudes below 87° on the morning side the 0.27 keV electron fluxes varied by three orders of magnitude over narrow latitudinal widths. On the morning side we note that the 2.7 keV electrons also underwent rapid fluctuations of two orders of magnitude. Count rates recorded during this quiet-time orbit in the 2.3 and 20 keV channels at latitudes below 68° are not due to auroral electrons but to energetic particles from the radiation belts penetrating the sensor casing.

The dayside cusp region appears in Figure 12-3 as a gap in the dayside discrete aurora near local noon. The cusp is a spatially limited region that extends in altitude out to the dayside magnetopause where magnetosheath plasma gains direct access to the magnetosphere. The intersection of the cusp with the polar ionosphere generally lies between 75° and 80° magnetic latitude in the noon sector. Consistent with their magnetosheath origin, electrons detected in the cusp region always have very soft spectra. In most cases cusp electrons have Maxwellian energy distributions with mean thermal energies less than 100 eV and essentially no electrons with energies greater than 1 keV. Again, it is expected that such electrons would produce ionization in the F-region ionosphere. Except when the cusp region is in darkness, these electrons will have little effect on height-integrated conductivities since solar EUV is the dominant ionization source in the E region where the conductivities peak.

Relatively uniform low energy electron fluxes into the polar cap have been characterized by Winningham and Heikkila (1974) as "polar rain". This type of precipitation can fill the entire polar cap. These particles have mean thermal energies  $\sim 100$  eV and are isotropically distributed in pitch angle outside of the atmospheric loss cone. The energy fluxes carried by these electrons range from  $10^{-2}$  to  $10^{-3}$  erg/cm<sup>2</sup>-sec, two or three orders of magnitude less than typical energy fluxes in the auroral oval. The similarity of the spectral slope of these particles to that of those measured in the dayside cusp suggests that "polar rain" electrons are of direct magnetosheath origin. Electron fluxes measured by Yaeger and Frank (1976) in the distant lobes of the magnetotail indicate that polar rain electrons enter the magnetosphere at great distances downstream from the Earth.

Embedded within the broader regions of "polar rain" are occasional enhanced fluxes of precipitating electrons. These electrons have mean thermal energies of  $\sim 100$  eV, but have undergone field-aligned accelerations that range from several tens of volts to approximately one kilovolt. Electrons that undergo accelerations of  $\sim 1$  keV are thought to be responsible for visible, Sun-aligned arcs in the polar cap. Those having undergone lesser accelerations are thought to be responsible for subvisual F-region aurora (Weber and Buchau, 1981). It should be noted that the enhanced polar cap fluxes associated with polar showers are most frequently observed during periods of northward IMF.

Localized, intense fluxes of electrons that have undergone field-aligned accelerations of several kilovolts can be found in the polar cap during periods of geomagnetic storms. Foster and Burrows (1976) observed that these fluxes, which appeared during the quieting period of a storm, were uniformly distributed over the polar cap. The ionization created by such electrons would be primarily at E-region altitudes.

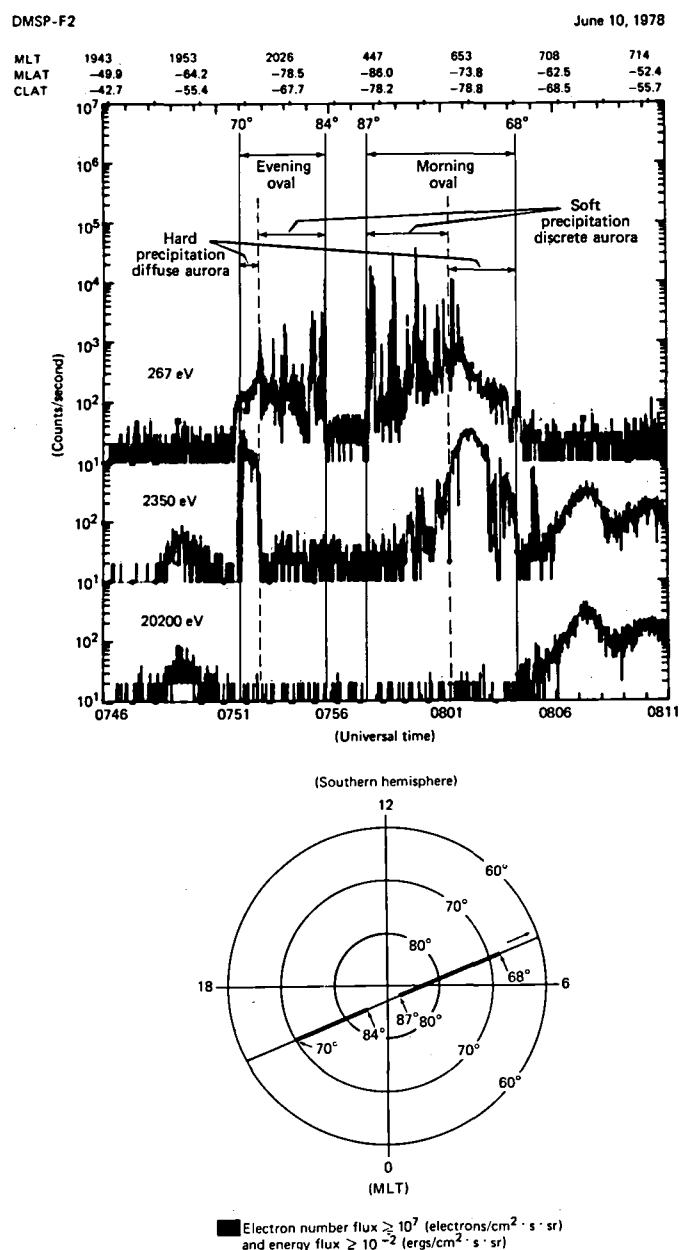


Figure 12-4. Example of polar electron precipitation profile obtained on June 10, 1978, during a very quiet condition, by a Defense Meteorological Satellite Program spacecraft (DMSP/F2). In the upper portion the simultaneous count rate profile of three selected channels (267 eV, 2.35 keV, 20.2 keV) as a function of universal time is depicted. Orbital information such as magnetic local time (MLT), corrected geomagnetic latitude (MLAT), and geographic latitude (CLAT) is given. The lower portion of the figure depicts the satellite trajectory in terms of magnetic latitude/magnetic local time.

For the most part the winter polar cap consists of regions of low height-integrated conductivities since the region is mostly in darkness and little precipitation reaches the E region. The summer polar cap can have a significant portion sunlit and as a result different conductivity characteristics. Occasionally energetic solar flare protons and electrons in the energy range greater than tens of keV penetrate directly into the polar cap and create significant ionization in the D region.

## ELECTRIC FIELDS

Characteristics of the local F-region ionospheric plasma motion can be examined by direct measurements of the thermal ion drift velocity or by direct measurement of the local electric field. In either case the deduced motion is the same.

When the interplanetary magnetic field has a southward component, there exists a large scale convection electric field that transports the plasma through the auroral oval towards local noon and towards midnight within the polar cap. The overall character of the pattern is variable and influenced by the solar ecliptic Y-component of the interplanetary magnetic field (Heppner, 1972). The characteristics of this two-cell convection pattern are dealt with in some detail by Working Group 11. We discuss its role in producing and redistributing plasma structures in Section IV.

Measurements of ion drift velocities and electric fields show that they possess considerable structure superimposed on the large scale sunward and antisunward convection. Figure 12-5 shows the typical signature that one observes in the high-latitude convection velocity when the IMF is southward. The polar diagram shows the location of the orbit track relative to a model two-cell convection pattern. In both the morning and evening sides of the high-latitude region the relatively smooth transition from corotation to significant velocities directed toward the Sun begins at the equatorial edge of the diffuse auroral precipitation. This transition may be equatorward of the diffuse optical emissions. The structure in the electric field throughout the diffuse auroral zone is quite pronounced and commonly changes in magnitude by a factor of 2 over spatial scales less than 100 km. In some cases changes of an order of magnitude at scale sizes ranging from several kilometers to several hundred meters can be observed. Perhaps the most important point to be emphasized, however, is that very little structure in the direction of the electric field can be expected in this region.

The transition from sunward to antisunward convection produces a natural reversal in the direction of the electric field. The reversal takes place over distances of 10-200 km and is associated with discrete precipitation events on the dusk side. Examination of the electric field within this region reveals, however, that a single, well defined reversal often does not exist. Rather, structures exist at scale sizes down to several hundred meters. Unlike the diffuse auroral zone these structures contain significant changes in electric field direction as well as magnitude and is therefore a region where strong velocity shears exist. The dawnside reversal shows similar characteristics in the electric field structure and differs only in that the average precipitating particle energy near the reversal is smaller than that on the evening side.

Electric field structure is not limited simply to the auroral zone but also exists in the cusp region and the polar cap. Data from DE-2 shown in Figure 12-6 indicate that structure in the cusp electric field exists at scale lengths from 1 km down to 1 m. This structure exists both in magnitude and direction although changes in direction of less than 45° are generally seen. Note the large magnitude variations at scale lengths of 1 to 10 km evident from Figure 12-6b.

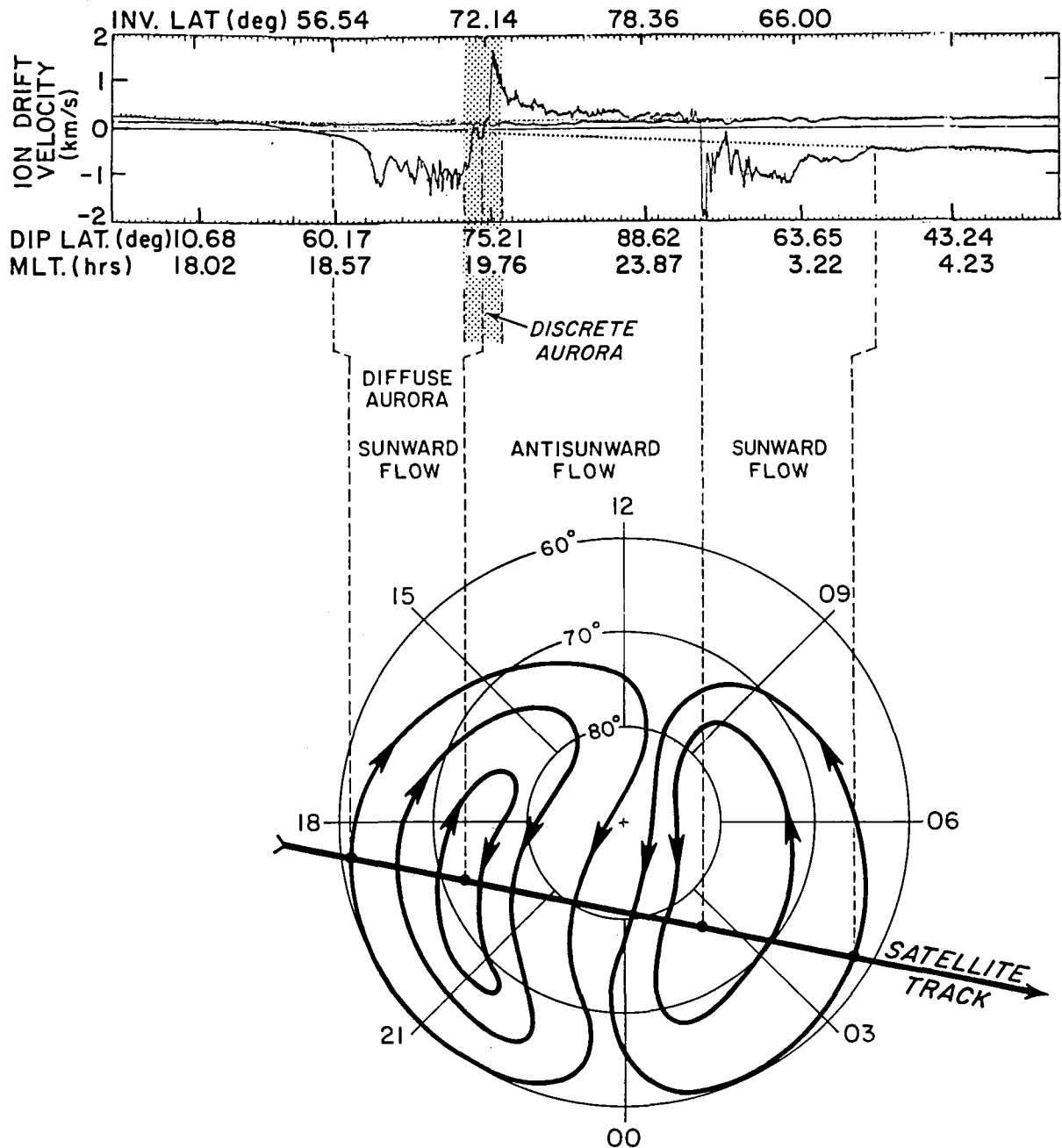


Figure 12-5. Typical signature of high latitude convection velocity with southward IMF. Ion drift velocity measured by AE-C during a pass across the high latitude Northern Hemisphere at an altitude of 350 km is exhibited in the upper portion. The lower portion shows the location of the satellite track relative to a model two cell convection pattern.

Larger-magnitude, rapidly-varying electric fields have been observed on auroral L-shells at high altitudes above the ionosphere. It has been postulated that the region above the ionosphere acts as an imperfect filter which limits the penetration of these structures (Mozer, 1981; Maynard et al., 1982).

In Figure 12-6 the strong broadband AC signals in the 4 Hz to 1 kHz range near 4:20 UT are commonly found in the region of the cusp. Curtis et al. (1982) concluded that these are local electrostatic manifestations of nonlocal field-aligned heating processes. Thus, they are signatures of the spatial variability of the field-aligned particle fluxes which in turn, ties the magnetic, electric, and density fluctuations together. Although this low frequency noise is common in the DE-2 data at all local times in the auroral oval, it is not always as strong.

During periods of southward IMF the large-scale electric field in the summer polar cap varies principally in magnitude. A wide range of spatial scales is again observed but the amplitude of the structure is generally smaller than that observed in the auroral zones. The most significant change taking place, when the interplanetary magnetic field has a northward component, is the frequent disappearance of a large-scale plasma circulation pattern. Under these conditions the highest latitudes show an increase in structure at spatial scales of 1 km and greater, and structures displaying changes in both magnitude and direction can be expected throughout the high latitude ionosphere. At shorter wavelengths structures are similar to those found in the auroral oval.

## FIELD-ALIGNED CURRENTS

Structure in the electric fields, magnetic fields, and precipitating particles are electrodynamically coupled. Large-scale field-aligned currents in the auroral oval are electrodynamically tied to the plasma convection. Their patterns, including the delineation of Region 1 and Region 2 currents and their general behavior, have been dealt with by Working Groups 7 and 11. Of particular interest to ionospheric structure studies is the fine structure in these currents. Currents in both directions varying over short spatial scales occur both in Regions 1 and 2. If the ionospheric conductivity is solar controlled then changes in  $\Delta B$  which are a measure of field-aligned currents, closely mirror the changes in  $E$  such as those shown in Figure 12-6b (Burke et al., 1982; Sigiura et al., 1982). The large rapid variations in  $\Delta B$  infer large field-aligned currents of the order of  $100 \mu A/m^2$  over scale sizes of a few km or less. These can be drivers for instabilities. The principal carriers of the upward currents are the precipitating electrons. Downward currents are presumed to be carried by upward thermal electron flows.

## PLASMA DENSITY STRUCTURE

Total plasma density and associated irregularity distributions provided important diagnostic parameters in understanding the multitude of active plasma processes at the equator. Similar data from a variety of sources can contribute substantially to an overall model of high-latitude ionization processes, irregularity scale size distributions, their transport, attendant instability mechanisms, and lifetimes. Some advances have already been made in these areas as a result of ground-based and in situ diagnostic systems.

In Figure 12-7 data from three passes of the S3-4 satellite across the high-latitude Northern Hemisphere show the variety of spatial scales and amplitudes of plasma density structure to be observed in the 200 km altitude region. The data for each orbit presents total in situ plasma densities,  $N_e$ , with the solar EUV ionization component displayed as the smoothly varying function  $N_{es}$ . Thus, the difference  $N_e - N_{es}$  represents density variations which are the result of

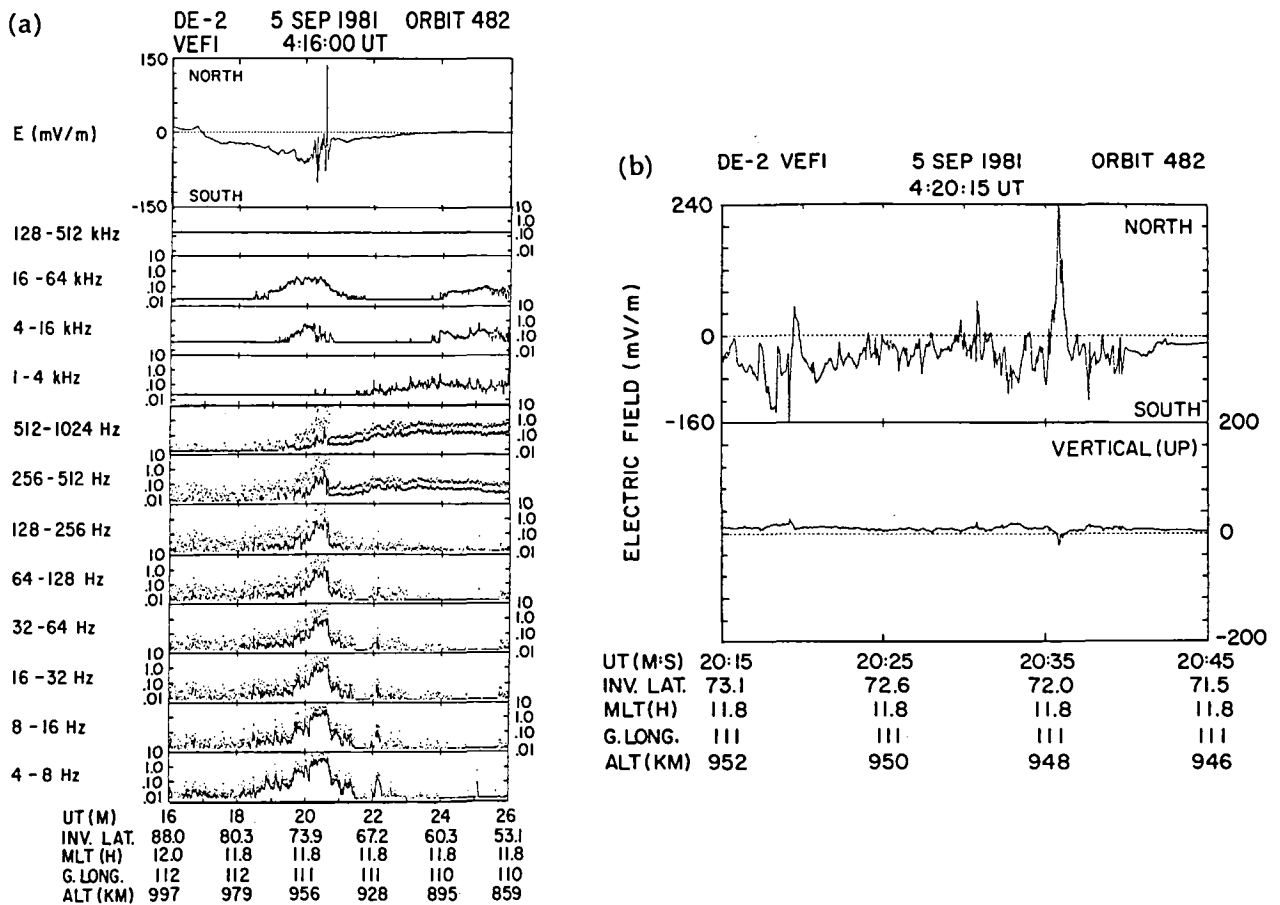


Figure 12-6. (a) Electric fields observed from a Southern Hemisphere polar cusp crossing of the DE-2 satellite on September 5, 1981. The DC data (upper panel) are 1/2 s averages and represent the electric field in the horizontal direction along the orbit track. The AC data are from a comb filter spectrometer. The AC measurement axis is at  $45^\circ$  to the horizontal and as near-perpendicular to  $\underline{B}$  as possible. The solid curve represents the RMS average value, while the dots are peak values (mV/M) (taken from Maynard et al., 1982).

(b) An expansion of the DE data in the region from 04:20:15 to 04:20:45 UT in (a). The time resolution is 16 sample/s, equivalent to 0.5 km. Note the scale change between the DC data in (a) and (b). The lower panel shows the vertical component of the field. Electric fields in the vertical direction should not be interpreted as field-aligned as this component can result from projections of the components perpendicular to  $\underline{B}$  to the north and east (taken from Maynard et al., 1982).

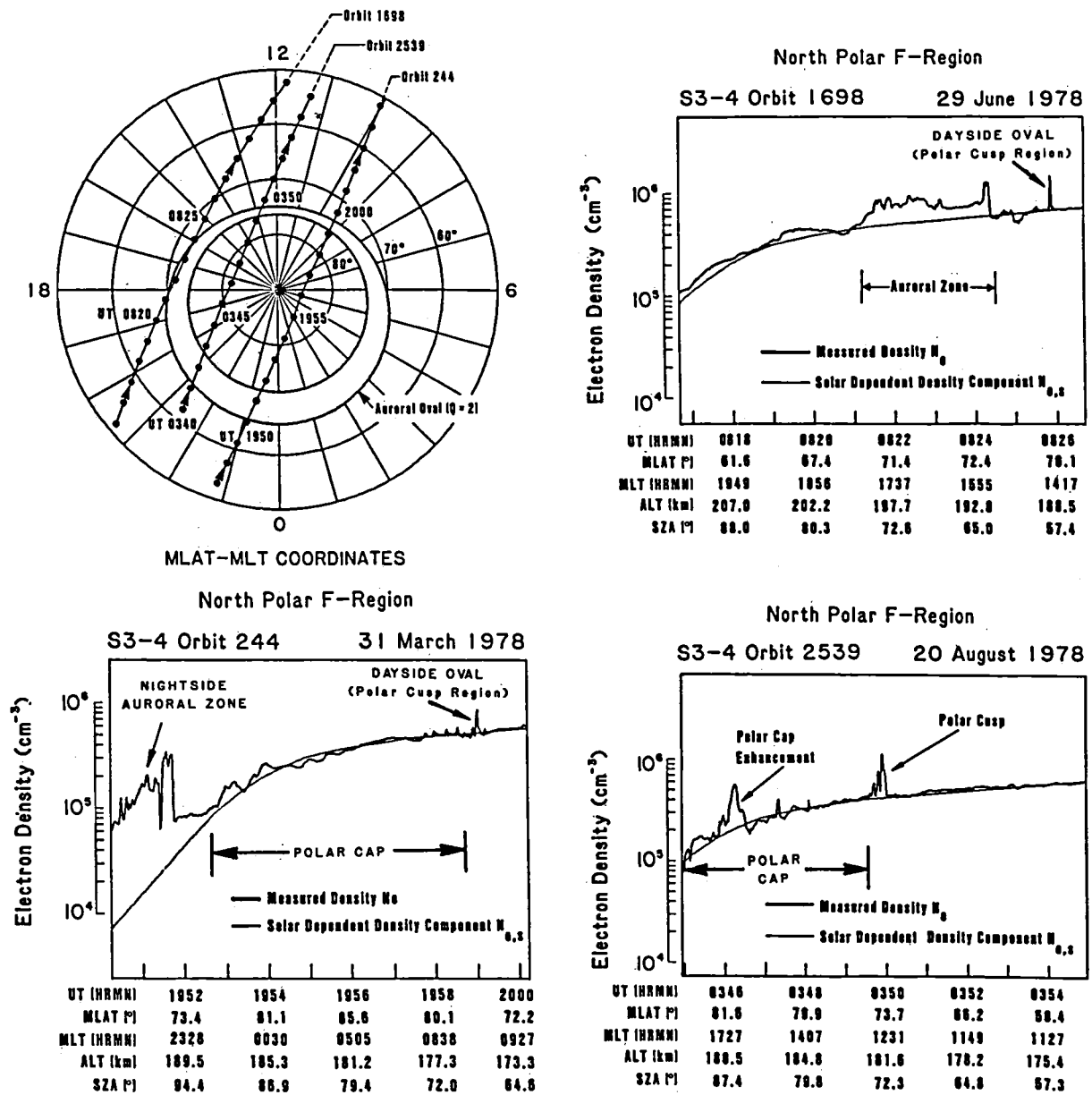


Figure 12-7. In situ plasma density structure across the northern high-latitude ionosphere as measured on three polar passes of the S3-4 satellite. Trajectories are shown in the upper left panel relative to a quiet auroral oval model. Total density  $N_e$  includes solar UV and precipitating-particle production. The difference  $(N_e - N_{e,s})$ , where  $N_{e,s}$  is the UV-source component, is due solely to energetic particle morphology, local plasma processes, and transport.

particle precipitation events, irregularity redistribution, ion chemistry, and convection. By referencing the data to the approximate position within the morphological domains of the high-latitude ionosphere (see upper left panel of Figure 12-7 and Figure 12-3) it is evident that both small and large scale structure always exist in the auroral oval and cusp region. In the overall S3-4 study (Szuszczewicz et al., 1982a, b; Rodriguez and Szuszczewicz, 1983) of northern-latitude irregularity structures ( $150 \text{ km} > \lambda > 20 \text{ m}$ ) observed during the Spring and Summer of 1978, it was clear that morphology and intensity distributions were identified with particle precipitation patterns in the auroral oval and the polar cusp region (more generally, the noon-sector auroral oval). The most intense density fluctuations observed under moderately active geomagnetic conditions ( $K_p < 2-3$ ,  $AE \sim 100$ ) occurred near the dayside oval in the region of the polar cusp, where smaller scale structure was particularly prominent and absolute density enhancements approached  $10^6 \text{ cm}^{-3}$ . Auroral oval irregularities displayed absolute enhancements generally less than those in the neighborhood of the polar cusp, with maximum values of about  $8 \times 10^5 \text{ cm}^{-3}$ . Large density enhancements within the polar cap were also detected principally toward the duskward side.

Efforts to characterize the spectral behavior of the structure at 200 km altitude indicate a  $k^{-2.5}$  behavior in discrete arcs, a softer  $k^{-1.8}$  dependence in the diffuse auroral zone and a return to a  $k^{-2.5}$  in the cusp, but with considerably greater variability in the spectral index than that observed in discrete arcs.

The obvious relationship between plasma density structures and the appearance of electron precipitation leads naturally to an altitude-dependent plasma-density profile representing the variability in incoming electron energy distributions. A perspective in this area is provided by Figure 12-8 showing contours of total ion concentration determined by the Chatanika radar as it scanned across the nightside auroral zone. Plasma density enhancements that are five times greater than the background ionosphere are observed here. On this occasion measurements of the local ion drift velocity and the electron temperature indicate that these enhancements were transported to the point of observation. In many other cases, however, similar plasma enhancements can be attributed to local particle precipitation that can also produce substantial E-region ionization at 120 km.

In the 300 km to 500 km altitude region convecting plasma enhancements such as those shown can be subject to instability mechanisms producing smaller scale irregularities that contribute to the spectral characteristics that are observed. Thus while the degree of variability in scale size and amplitudes to be observed at 500 km is as great as that seen at 200 km, the evolution of the structures and the mechanisms that change the spectral characteristics may be quite different.

Note that we have used only a small subset of data to illustrate the degree of structure in the high-latitude ionosphere. Additional data produced by a large variety of in situ and remote sensing techniques include radio wave scintillation measurements (Aarons, 1973; Fremouw et al., 1977; Buchau et al., 1978; Rino et al., 1978), satellite measurements (McClure and Hanson, 1973; Dyson et al., 1974; Sagalyn et al., 1974; Phelps and Sagalyn, 1976; Clark and Raitt, 1976), rocket observations (Kelley et al., 1980), ground-based optics (Weber and Buchau, 1981) and radars (Vickery et al., 1980; Greenwald et al., 1978; Hanuise et al., 1981). The smallest scale size at which plasma density structure is observed depends critically on the measurement techniques involved, and one must carefully consider the limitations of instruments in comparing results.



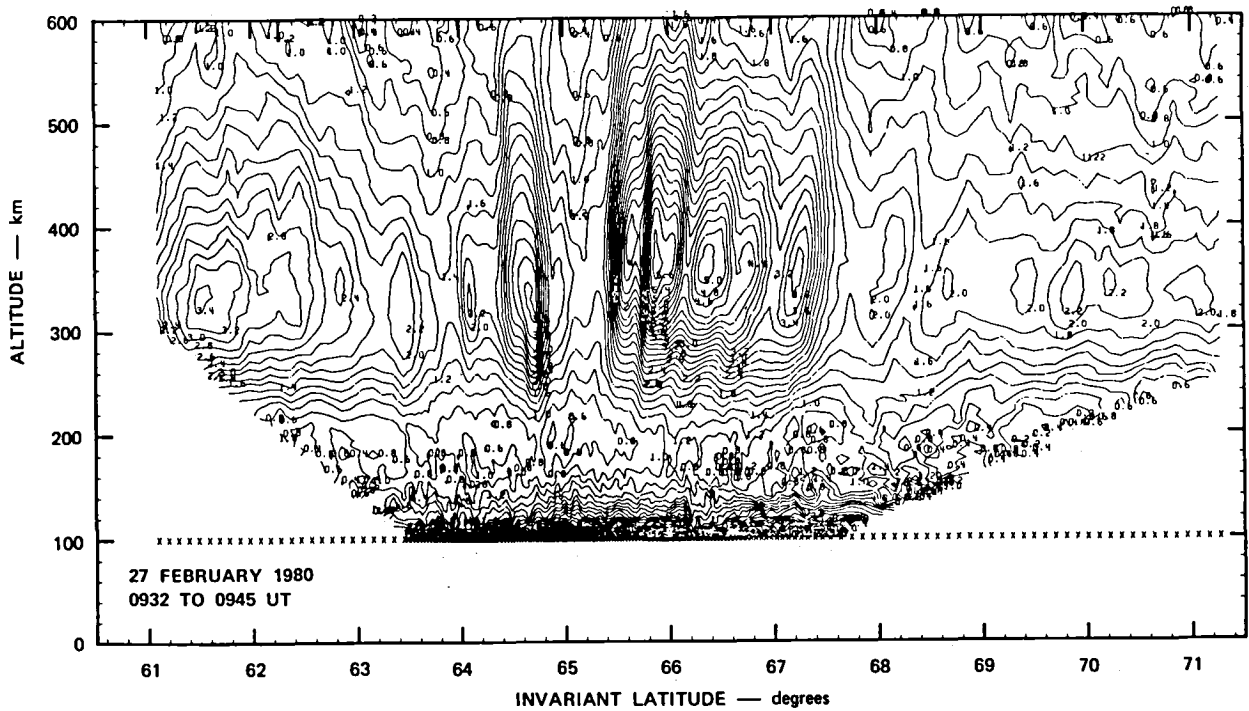


Figure 12-8. Altitude/latitude variation of electron density in the midnight-sector auroral zone measured by the Chatanika radar (taken from Kelley et al., 1982c).

### III. PLASMA INSTABILITY THEORY

Because of the diverse scale sizes associated with ionospheric irregularities, theoretical studies are usually broken into two parts: macrophenomena and microphenomena. We define "micro" to denote fluctuations with wavenumbers satisfying  $ka_i \gg 1$ , where  $a_i$  is a thermal ion gyroradius. Although our theoretical understanding of high-latitude ionospheric irregularities is not on a par with that of equatorial ionospheric irregularities, much equatorial instability theory is applicable to the high latitude environment. In this section we discuss both the macroinstability and microinstability regimes.

#### MACROINSTABILITIES AND HIGH LATITUDE STRUCTURE

In this subsection we discuss the contributions of plasma macroinstabilities to the nature of large scale high-latitude ionospheric structures.

##### F-Region Irregularities

In the plasma fluid approximation the equations describing the basic dynamics of the electron and ion high-latitude F-region plasma are the electron and ion continuity and momentum equations and  $\nabla \cdot \underline{J} = 0$  where  $\underline{J}$  is the current (Keskinen and Ossakow, 1982). Since the fluid picture corresponds to low frequencies and long wavelengths, the inertial terms in the electron and ion momentum equations are ignored. The assumption of quasi-neutral fluctuations  $n_e \approx n_i = n$  implies  $\nabla \cdot \underline{J} = 0$ , where  $n$  denotes density and the subscript e(i) denotes electron (ion).

The following geometry is used: the y axis, is parallel to  $\nabla n_0$  and in the north-south direction, the x-axis is parallel to  $\underline{E}_0$  the ambient electric field pointing west, and the z-axis is downward along the magnetic field,  $\underline{B}$ . Linearizing the equations and assuming that the fluctuating density and potential are proportional to  $\exp [i(k_x x + k_z z - \omega t)]$ , with  $\omega = \omega_r + i\gamma$  and  $L^{-1} \equiv (1/n_0)(\partial n_0 / \partial y)$ , yields a growth rate

$$\gamma = \left[ -\frac{\nu_{ei}}{\Omega_e} \frac{1}{L} \left( \frac{\nu_{in}}{\Omega_i} \frac{cE_0}{B} - \theta V_d \right) \right] \quad (1)$$

$$\div \left( \theta^2 + \frac{\nu_{in}}{\Omega_i} \frac{\nu_{ei}}{\Omega_e} \right) - D_{\perp} k_x^2 - D_{\parallel} k_z^2$$

where

$$\theta \equiv k_z/k_x, \quad \underline{V}_d = \hat{z}(v_0 - V_0), \quad k_{\perp}^2 = k_x^2, \quad k_{\perp}^2 \gg k_z^2, \quad D_{\perp} = (\nu_{ei}/\Omega_e \Omega_i) c_s^2$$

and

$$D_{\parallel} = (c_s^2/\nu_{in}) [1 + [(\nu_{in}/\Omega_i)^2 / ((\nu_{ei}\nu_{in}/\Omega_e \Omega_i) + (k_z^2/k_{\perp}^2))]$$

( $\parallel$  denotes along z). Here  $v_0$  and  $V_0$  refer to the electron and ion velocities along the magnetic field giving rise to a cold magnetic field aligned current. The symbol  $\nu_{in}$  denotes the ion-neutral collision frequency,  $\nu_{ei}$  the electron-ion collision frequency,  $c$  the speed of light,  $T_e \approx T_i \equiv T$  the species temperature,  $c_s$  the ion acoustic speed, and  $\Omega_i(\Omega_e)$  the ion (electron) gyrofrequency. We have neglected  $\nu_{en}$  compared with  $\nu_{ei}$  and taken  $\nu_{\alpha}/\Omega_{\alpha} \ll 1$  for  $\alpha = i, e$  (F-region approximation).

The expression for the growth rate  $\gamma$  in (1) can be maximized as a function of  $\theta = k_z/k_x$ , a measure of field alignment, using  $\partial\gamma/\partial\theta|_{\theta=\theta_m} = 0$  giving:

$$\theta_m = \frac{v_{in}}{\Omega_i} \frac{cE_{ox}}{B V_d} \pm \left[ \left( \frac{cE_{ox}}{B V_d} \right)^2 \left( \frac{v_{in}}{\Omega_i} \right)^2 + \left( \frac{v_{ei} v_{in}}{\Omega_e \Omega_i} \right) \right]^{1/2} \quad (2)$$

Using typical diffuse auroral F-region parameters

$$v_{in}/\Omega_i \approx 10^{-4}, v_{e}/\Omega_e \approx 10^{-4}, E_{ox} \approx 10 \text{ mV/m}, j_{||} = n_0 e V_d \approx 1 \mu\text{A/m}^2, B = 0.5\text{G}, n_0 \approx 10^5 \text{ cm}^{-3},$$

this gives  $|\theta_m| \approx 10^{-4}$ , i.e. approximate field alignment. Inserting these parameters into (1) with  $L \approx 20 \text{ km}$ ,  $D_{\perp} \approx 0.2 \text{ m}^2/\text{s}$ , and  $D_{||} \approx 10^8 \text{ m}^2/\text{s}$ , we find that the fastest growing linear modes have growth times  $\gamma_{\max}^{-1} \approx 10^2 \text{ s}$ .

It is convenient to discuss Equation (1) in two limits. If  $k_z = 0$ , then  $\gamma \approx cE_0/B L$  (ExB gradient drift instability) while if  $E_0 = 0$ ,  $\gamma_{\max} \approx (V_d/2L) (1 + v_{in}\Omega_e/\Omega_i v_e)^{-1/2}$  corresponding to the current convective instability. We now discuss these two instabilities.

It is well known that convecting ionospheric plasma clouds are unstable, under certain conditions, and can produce plasma structure through the ExB gradient drift instability. This convective instability (Simon, 1963; Linson and Workman, 1970) has a nonlinear evolution (Zabusky et al., 1973; Scannapieco et al., 1976) resembling the classical Rayleigh-Taylor instability described in Section 1. The basic ExB instability mechanism can be understood by noting Figure 12-9. The ambient background electric field  $E_0$  is in the  $-x$ -direction, the ambient magnetic field  $B_0$  in the  $z$ -direction, and the background density gradient points in the  $y$ -direction. Let the density be perturbed by a small amplitude sinusoidal perturbation with wavenumber  $k$  parallel to  $E_0$ . In the F region, the ions drift to the left in the Pedersen direction relative to the electrons. This gives rise to space charges (+ and -) which in turn cause small scale electric fields  $E'$  alternating in direction as shown. The corresponding  $E' \times B_0$  drifts will then convect depleted regions upward (toward increasing density) and enhanced regions downward (toward decreasing density) with the result that they both appear to grow relative to the background density gradient--an unstable situation. In the previous configuration the ExB gradient drift instability can also arise with a neutral wind  $U$  blowing in the  $-y$  direction with no  $E_0$ .

The presence of the highly conducting auroral E layer may reduce  $E'$  and hence the growth rate by the ratio of the E- and F-region height-integrated Pedersen conductivities. This reduction factor can be as high as 100 for typical auroral conditions. It should be noted that Equation (1) does not contain the effects of E region conductivity.

Using both analytical and numerical simulation techniques, Keskinen and Ossakow (1982) have studied the linear stability and nonlinear evolution of large scale convecting plasma enhancements in the auroral F-region ionosphere. Their results indicate that convecting diffuse auroral plasma enhancements can be driven unstable through the ExB gradient drift instability, generating plasma density and electric field structure with scale sizes of tens of kilometers to tens of meters. These irregularities take the form of striation-like structures, elongated in the north-south direction for equatorward convection, which can form on the order of half an hour in the absence of an auroral E layer. Figure 12-2 illustrates the model of the plasma enhancement used in the numerical simulations. In a plane nearly perpendicular to the geomagnetic field, Figure 12-10 displays the linear and nonlinear evolution of the ExB instability in a convecting auroral F-region plasma enhancement. Parameters were used that agree with the observations of

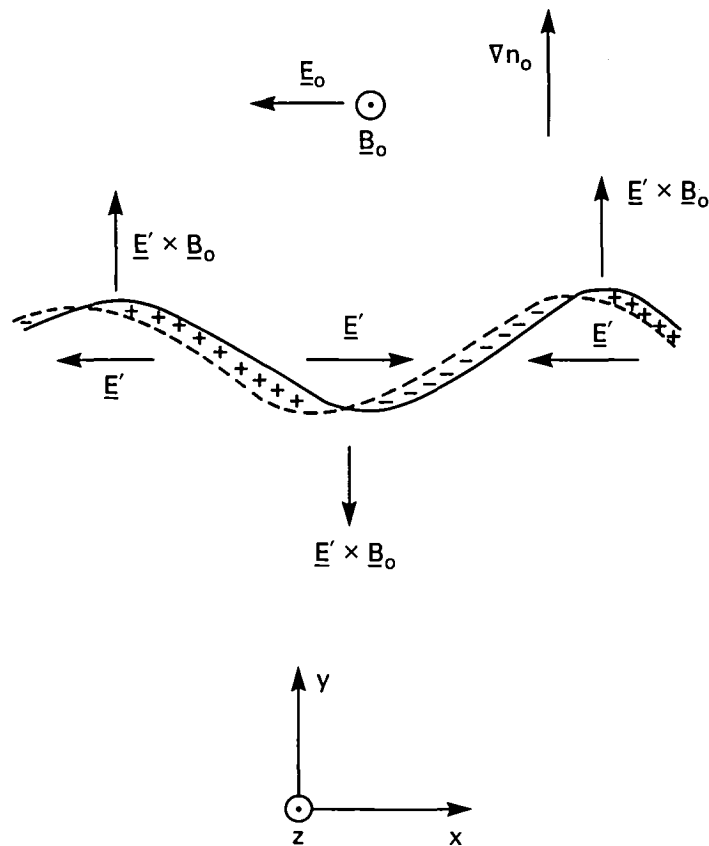


Figure 12-9. Physical mechanism of the F region  $\underline{E} \times \underline{B}$  gradient drift instability.

Vickrey et al., (1980), including field-aligned currents. Figure 12-10a shows the initial configuration of the ionization enhancement and includes a small 1% density perturbation. Figure 12-10b illustrates the linear regime at  $t=550$  sec and shows unstable growth on the poleward side of the equatorward convecting plasma enhancement. Figure 12-10c gives the structure of the plasma enhancement at  $t=1000$  sec. Finally, Figure 12-10d displays the plasma enhancement at  $t=1600$  sec in the well-developed nonlinear regime where steepened and elongated striation-like structures are evident. These fingers or striations will be oriented in a direction dependent upon the local electric field magnitude and direction.

These theoretical results are not inconsistent with recent experimental observations, using the Chatanika radar, which indicate large neutral wind velocities ( $\underline{E} \times \underline{B}$  gradient drift instability) and fingerlike structures collocated with large scale convecting plasma enhancements (see Figure 12-11). Vickrey and Kelley (1982) have studied the role of classical diffusion and a conducting E layer in removing these irregularities once they are produced. In addition, Keskinen and Ossakow (1982) found that the large scale size irregularities (fingers) can cascade to smaller scale size structures through nonlinear mode coupling. The spatial power spectra of these irregularities in the north-south  $P(k_y)$  and east-west  $P(k_x)$  direction taken in the nonlinear regime of the simulations can be well represented by power laws of the form  $P(k_x) \propto k_x^{-n_x}$  with  $n_x \approx 2-2.5$  for  $2\pi/k_x$

between 100 km and 3 km while in the north-south direction  $P(k_y): V k_y^{-n_y}$  with  $n_y \approx 2$  for  $2\pi/k_y$  between 256 km and 3 km. This process of finger formation, elongation, and steepening is almost self-similar in character with similar morphologies and power spectra for scale sizes  $\lambda$  between 1 km and 100m (Keskinen and Ossakow, 1983). Some observational evidence (Vickrey et al., 1980) indicates that these plasma enhancements are probably subjected to ambient auroral convection patterns. As a result, these enhancements could be a major source of F-region ionospheric structure throughout the auroral zone and polar cap (Kelley et al., 1982c).

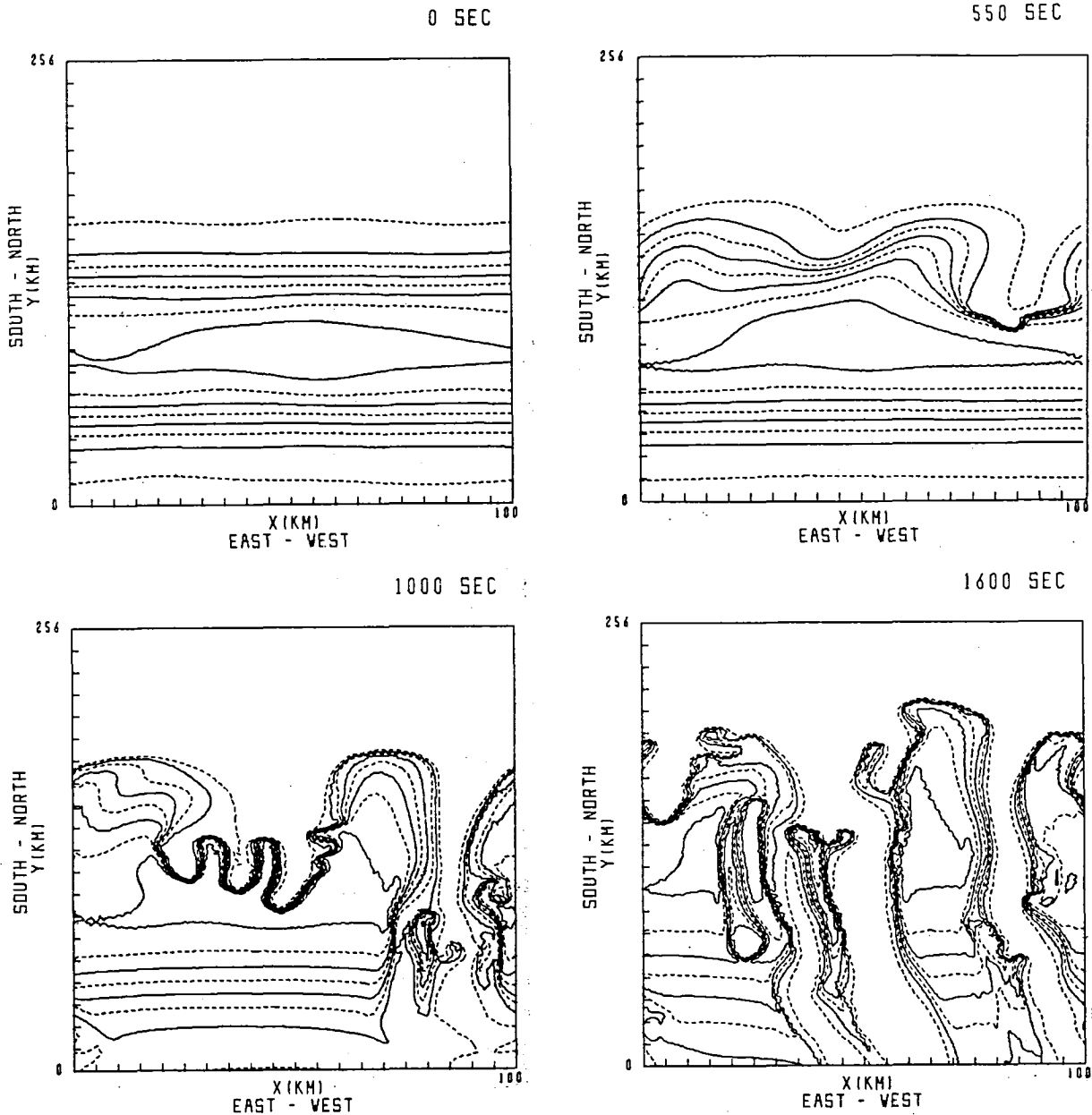


Figure 12-10. Real space plasma isodensity contour plots of  $[(n_o(x,y) + \delta n(x,y))/N_o]$  at (a)  $t=0$  sec, (b)  $t=550$  sec, (c)  $t=1000$  sec, (d)  $t=1600$  sec. The y axis is compressed by a factor of 2.58. The distance between tick marks in the x direction (y direction) is 5 km (12.8 km). Eight contours are plotted in equal increments of 1.25 beginning at 1.25. The observer is looking upward along the magnetic field lines (taken from Keskinen and Ossakow, 1982).

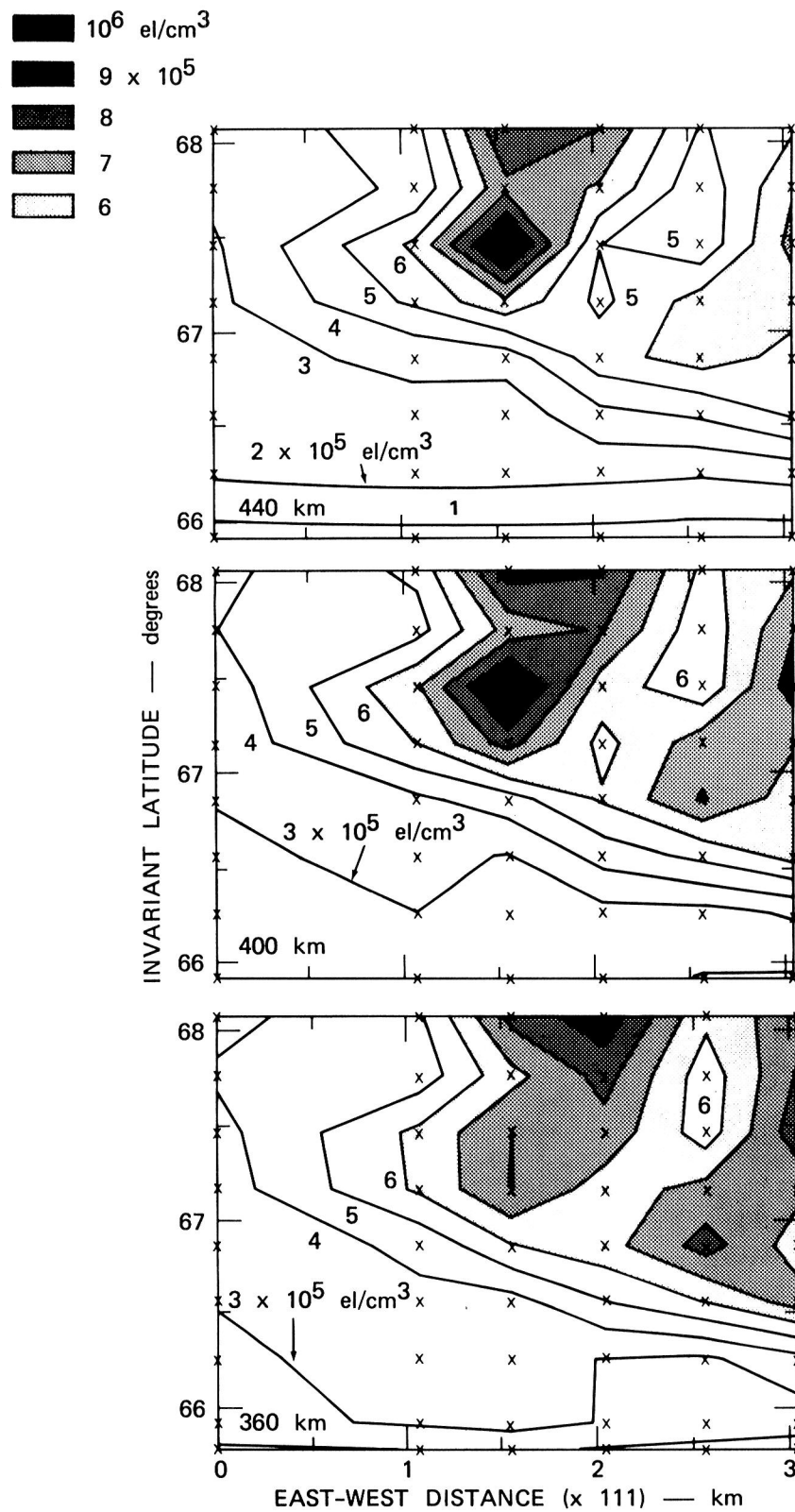


Figure 12-11. East-west structure/latitude variation of electron density at three different altitudes on 16 November, 1981.

The coupling of density gradients and magnetic field aligned currents can also lead to plasma instability through the current convective instability (Lehnert, 1958; Kadomtsev and Nedospasov, 1960). Ossakow and Chaturvedi (1979) showed that plasma enhancements can also be linearly unstable to the current convective instability in regions where the  $\underline{ExB}$  gradient drift instability is stable, i.e., in regions where the local convection velocity and density gradient are in opposite directions. The basic physical mechanism responsible for the current convective instability, which is similar to the  $\underline{ExB}$  gradient drift instability, is as follows (see Figure 12-12): Let the magnetic field  $\underline{B}_0$  and current  $j_{0\parallel}$  be in the z-direction, and ambient electric field  $\underline{E}_0$  with perpendicular component in the x-direction, and a density gradient in the y-direction. Consider a density perturbation with wavevector  $\underline{k}$  as shown. The projection on  $\underline{k}$  of the ion Pedersen drift caused by  $\underline{E}_0$  results in a drift that is stable to the  $\underline{ExB}$  instability. However, the assumed direction of  $j_{0\parallel}$  implies a relative drift between ions and electrons (in the electron rest frame) that is antiparallel to  $\underline{B}_0$ . This motion projected on  $\underline{k}$  gives space charges and subsequent electric fields or  $\underline{E}'$ . If the particle motion projected on  $\underline{k}$  is dominated by  $j_{0\parallel}$  then the total perturbation space charge electric fields will be denoted by  $\underline{E}'$ . The corresponding  $\underline{E}' \times \underline{B}_0$  drifts will then convect enhanced (depleted) regions out of (into) the figure which will appear to grow relative to the background density in direct analog to the  $\underline{ExB}$  gradient drift instability as outlined previously. The ratio of the linear growth rate of the  $\underline{ExB}$  gradient drift instability versus the current convective instability can be written approximately

$$(2V_0/V_d)(1 + \Omega_e v_i / \Omega_i v_e)^{1/2} \approx 2V_0/V_d \approx 2V_0 n_e / j_{0\parallel}$$

since  $\Omega_e v_i / \Omega_i v_e$  is of order unity in the auroral F region. Here  $V_0 = cE_0/B$  is the convection speed across the magnetic field. As a result, for strong perpendicular electric fields and/or weak currents and ambient densities the  $\underline{ExB}$  gradient drift instability will dominate and vice versa. Also, as in the previous case, an auroral E layer may slow the growth of this instability.

Using plasma fluid equations to describe the diffuse auroral F region, Ossakow and Chaturvedi (1979) showed that by taking account of the diffuse auroral particle precipitation (current), and the plasma density gradient (Rino et al., 1978), this stable  $\underline{ExB}$  gradient drift instability geometry becomes unstable to the current convective instability. A maximum linear growth rate was derived and favored creation of the plasma instability at an altitude of 400 km over say 200 km. Although a finite wavevector parallel to the magnetic field is required for the unstable wave, the instability is mainly magnetic field aligned. This instability is fluid-like in nature and so could directly account for the long wavelength diffuse auroral scintillation causing irregularities as can the  $\underline{ExB}$  instability. This simple model appeared to explain most of the Wideband satellite observational features reported by Rino et al. (1978).

The above linear theories could not explain the apparent L shell aligned nature of the observed F-region irregularities. Chaturvedi and Ossakow (1979) have shown that, using a nonlinear theoretical mode coupling analysis with two modes, the current convective instability can stabilize nonlinearly by generating linearly damped harmonics along the density gradient. An estimate of the saturated amplitudes showed good agreement with some recent rocket data. With a northward plasma density gradient, the dominant nonlinear harmonic is in the northward direction and so is in agreement with Wideband satellite observations regarding the direction of these F-region irregularities. Also, the power spectrum for these dominant harmonics is  $\propto k^{-2}$ . The nonlinear evolution of the current convective instability in the auroral F-region ionosphere was investigated using many modes by Keskinen et al. (1980). The evolution and morphology of the plasma density irregularities generated by the current convective instability was similar to that of the  $\underline{ExB}$  instability. In addition, the linear theory of the current convective instability has

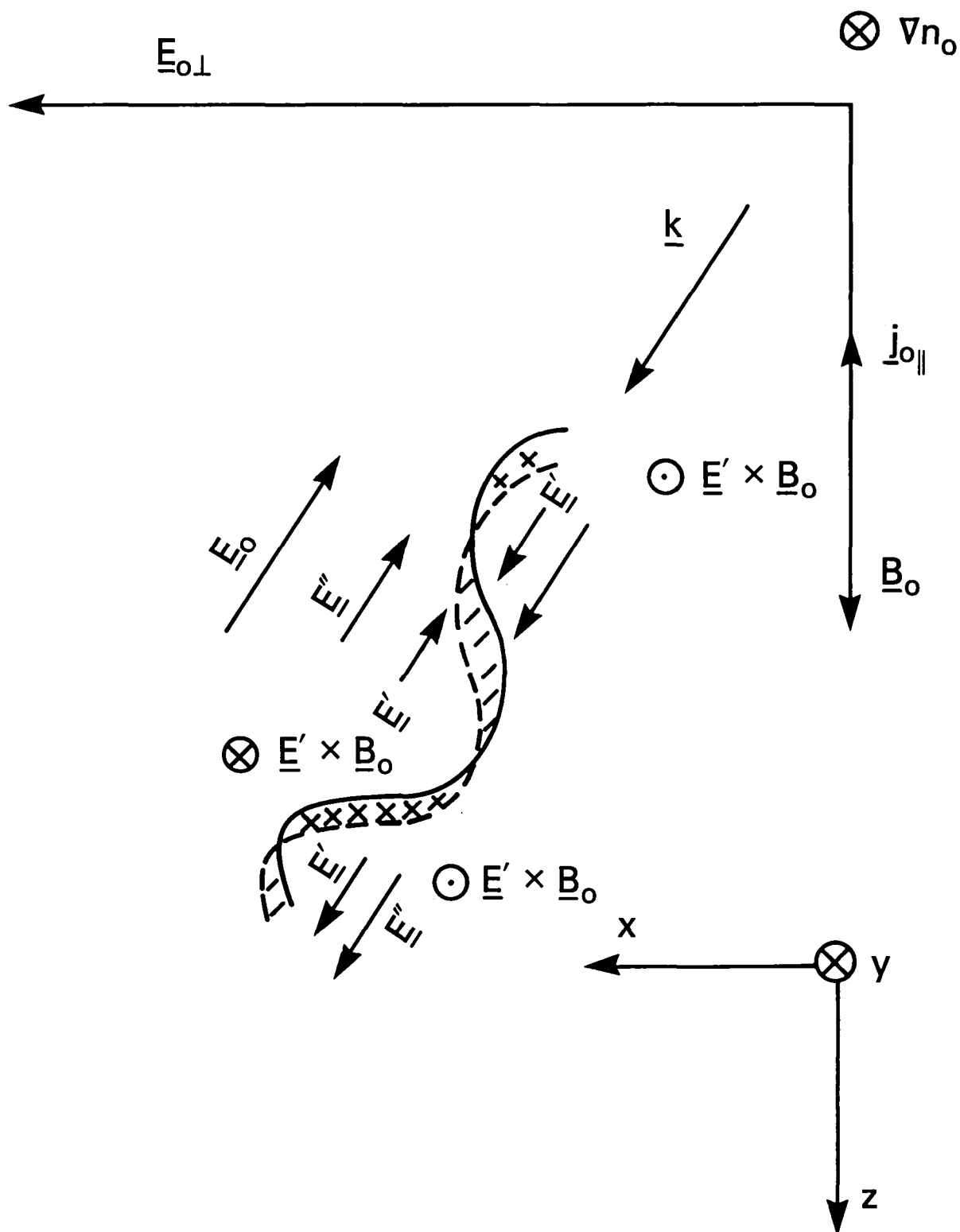


Figure 12-12. Physical mechanism of the current convective instability (taken from Ossakow and Chaturvedi, 1979).



been extended to include magnetic shear effects (Huba and Ossakow, 1980); highly collisional (non-magnetized) ions, electromagnetic effects, and ion inertial effects, which extends the instability to higher F-region altitudes (Chaturvedi and Ossakow, 1981); and the effects of a warm electron beam (Chaturvedi and Ossakow, 1983), which prove to be negligible.

Another source of plasma free energy that could directly excite small and large scale irregularities in the high latitude F-region ionosphere is velocity sheared plasma flows, e.g., near auroral arcs (Kelley and Carlson, 1977). For unstable velocity sheared plasma flows perpendicular to a magnetic field, the growth is nonlocal and maximizes for irregularity wavenumbers  $k$  such that  $kL_v \lesssim 1$  (Mikhailovskii, 1974) where  $L_v$  is the velocity shear gradient scale length. A local instability ( $kL_v > 1$ ) can operate for velocity sheared plasma flow parallel to a magnetic field (Mikhailovskii, 1974).

Still another source of free energy in the high-latitude F-region ionosphere is the magnetic field-aligned current system. Chaturvedi (1976), by invoking the collisional ion cyclotron instability, has shown that strong field-aligned currents in the auroral F-region ionosphere can lead to irregularities with scale sizes of hundreds of meters perpendicular to the magnetic field ( $k_{\perp} a_i \ll k_{\parallel} a_i < 1$ ). Relative electron-ion velocities of several kilometers per second are needed to excite this instability.

### E-Region Irregularities

In the high-latitude E region, experimental observations using both radar backscatter and rockets have shown the existence of plasma density irregularities with scale sizes of meters to hundreds of meters. These irregularities have been shown to be highly correlated with the auroral electrojet. Various plasma instabilities and processes similar to those operating in the equatorial electrojet have been successful in accounting for many of the features of these irregularities. Fluid equations that can describe the basic dynamics of the auroral E-region electrojet plasma are the electron and ion continuity and momentum equations and quasi-neutrality (Rogister and D'Angelo, 1970; Sudan et al., 1973). Here, as in the F region, electron inertia is ignored; in contrast to the F-region case, however, ion inertia is retained and the ions are taken to be collision-dominated and thus unmagnetized.

Linearizing these equations and assuming a relative drift  $\underline{V}_d$  between electrons and ions across the magnetic field, we find the frequency and growth rate of waves propagating perpendicular to the magnetic field:

$$\omega_{\underline{k}} = \underline{k} \cdot \underline{V}_d / (1 + \psi) \quad (3)$$

$$\gamma_{\underline{k}} = (\psi / (1 + \psi)) [(\Omega_e / \nu_e)(\omega_{\underline{k}} / kL) + (\omega_{\underline{k}}^2 - k^2 c_s^2) / \nu_i] \quad (4)$$

where

$$\psi = \nu_i \nu_e / \Omega_i \Omega_e, \quad L = |n_0 / \nabla n_0|$$

with  $\nabla n_0$  the electron density gradient along  $\underline{E}_0$ , and  $c_s = [(K(T_e + T_i) / m_i)]^{1/2}$  the ion acoustic velocity. It is convenient to discuss two well known limits. If  $L \rightarrow \infty$  (no density gradient) then (3) and (4) reduce to the Farley-Buneman instability (Farley, 1963; Buneman, 1963) with instability threshold condition  $V_d > c_s$ . If  $0 < V_d \ll c_s$  and the density gradient is retained such that  $(\Omega_e / \nu_e) \cdot (\nu_i / \omega_k) (1 / kL) > 1$ , then the  $\omega_k^2$  term in (4) can be neglected recovering the  $\underline{E} \times \underline{B}$  gradient drift instability with approximate instability criterion  $(\Omega_e / \nu_e)(V_d / L) > (k^2 c_s^2 / \nu_i)$ , when  $\underline{k}$  is along the  $y$  direction.

High latitude E-region structure with wavelengths up to several hundreds of meters have been identified principally through in situ rocket probes. Vertical density gradients are often observed in conjunction with auroral electrojet currents. As a result, the ExB gradient drift instability has been invoked to explain this large scale structure. This instability is somewhat different from the ExB instability used to describe high latitude F-region irregularities (see Section III entitled "F-Region Irregularities"). In the F region, the basic current flow is in the Pedersen direction, while the Hall current dominates in the E region. However, the physical mechanisms describing the ExB instability in both regions are essentially identical. Again, the linear theory of the ExB instability cannot account for irregularities propagating perpendicular to the electrojet current and saturated wave amplitudes and spectra. Several nonlinear theories of the E-region gradient drift instability, e.g., quasi-linear effects (Sato and Ogawa, 1976), resonance broadening (Weinstock and Williams, 1971) and strongly turbulent nonlinear wave-wave interactions (Sudan and Keskinen, 1977) have been proposed.

## MICROINSTABILITIES AND HIGH LATITUDE STRUCTURE

### General Comments

Here we review the existing theory of plasma microinstabilities that may contribute to the evolution of long wavelength structures in the high latitude ionosphere below 1000 km. We consider the following framework: various sources of plasma free energy drive plasma microinstabilities, which may be described adequately by linear theory during their growth phase (see next section on "Linear Theories"). These instabilities lead to enhanced fluctuations which, in turn, scatter the plasma particles leading to cross-field diffusion and other forms of transport; these processes require a nonlinear description (see Section on "Nonlinear Effects"). These transport expressions may then be incorporated into fluid codes so that their effects on the evolution, and in particular the dissipation, of macroscopic structure may be determined.

### Linear Theories

One begins with a linearized kinetic equation (usually the Vlasov equation with an appropriate collision term) and Poisson's equation. The derivation and detailed form of the kinetic dispersion equation depend upon the zeroth order plasma properties and especially on the type of free energy under consideration. We do not discuss such details here, but merely cite appropriate references. We now categorize various instabilities by their sources of free energy.

**1.  $\nabla n \perp B$ :** The first source of free energy we consider is a density gradient perpendicular to  $B$  which can arise in the ionosphere, for example, due to the long wavelength ExB gradient drift instability. An important parameter for characterizing the resultant density drift microinstabilities is  $\epsilon_n \equiv (1/n)(dn/dx)$ , the inverse density gradient scale length. We assume weak density gradients, i.e.,  $\epsilon_n a_i \ll 1$ .

The universal density drift instability is driven by the electron diamagnetic drift  $v_{ne}$  ( $v_{ne} = -\epsilon_n v_e^2 / 2\Omega_e$ , where  $v_e$  is the electron thermal velocity and  $\Omega_e$  is the electron gyrofrequency). This permits a drift wave,  $\omega \approx k_y v_{ne}$ ,  $0 < k_z \ll k_y$ , to exist in the plasma. This mode can arise in density gradients that are very broad compared to an ion gyroradius, so  $v_{ne} \ll v_i$  (ion thermal speed); since for this instability  $k_y a_i \lesssim 1$ , it follows that  $\omega_r \ll \Omega_i$  (ion cyclotron frequency). In addition,  $\gamma \lesssim \Omega_i$ . Ionospheric applications of the universal instability in the collisionless limit have been discussed by Costa and Kelley (1978a, b), Thomsen and Gary (1982), and Bernhardt et al. (1983). The role of ionospheric collisions on this instability has been treated by Huba and Ossakow (1979), and Gary et al. (1983).

Another class of density drift instabilities is driven by the ion diamagnetic drift velocity,  $v_{ni}$  ( $v_{ni} = \epsilon_n v_i^2 / 2\Omega_i$ ). At small  $k_y$  the associated drift wave satisfies  $\omega_r \approx k_y v_{ni}$ , but at  $k_y a_i \gg 1$ , the mode has substantial dispersion, so that at small  $\epsilon_n a_i$ ,  $\omega_r < \Omega_i$ . At perpendicular propagation there is also a Bernstein or ion cyclotron wave at  $\omega_r > \Omega_i$ . As  $\epsilon_n a_i$  is increased, the frequency of the drift wave increases until it reaches the ion cyclotron frequency; at this point an ion cyclotron instability can arise.

As the drift speed is further increased, the growth rate of this instability increases until  $\gamma \approx \Omega_i$ . In this case one can show that the ion response to the fluctuating fields becomes unmagnetized. In this regime the mode is called the lower-hybrid density drift instability. In effect it is driven unstable by inverse Landau damping on the positive slope of the  $v_y$ -dependent part of the ion distribution function. By comparison to the universal density drift instability, somewhat steeper gradients are required to excite the lower-hybrid density drift instability in the ionosphere. Huba et al. (1978) first pointed out that the lower-hybrid density drift instability might arise in the ionospheric F region; the detailed effects of collisions on this mode have also been studied by Sperling and Goldman (1980) and Huba and Ossakow (1981a). Threshold plots for this instability in the equatorial F-region are given by Huba and Ossakow, (1981b) and Gary et al. (1983).

Gary et al. (1983) have compared these two instabilities and have shown that, in general, the universal is the density drift instability most likely to be driven unstable by density gradients at  $\epsilon_n a_i \lesssim 0.05$  in the equatorial ionosphere. Using typical parameters, they showed that this mode may have appreciable growth at about 300 km and above. This result should, in general, carry over to the auroral ionosphere with the additional constraint that a shorter wavelength component parallel to  $\underline{B}$  is required, allowing the universal to grow only at somewhat steeper density gradients.

**2.  $\underline{J} \parallel \underline{B}$ :** The next source of free energy we consider is an electron drift velocity  $\underline{v}_{oe}$  parallel to  $\underline{B}$ . In the context of the auroral ionosphere, this corresponds to a field-aligned current ( $I_z$ ) carried by the entire electron distribution function. The origin of such currents has recently been reviewed by Stern (1983). In a collisionless plasma, this source of free energy leads to two unstable electrostatic modes, the ion acoustic and the ion cyclotron current instability. Under typical ionospheric conditions, Kindel and Kennel (1971) have shown the latter has the lower threshold. However, typical auroral ionospheric conditions below 1000 km lie far below the threshold of this instability, and this mode is not considered to make a significant contribution in this region.

Now consider a cross-field density gradient as well as a field-aligned current. Given these sources of free energy in a collisional plasma, Kadomtsev (1965) used a fluid model to derive the properties of the current convective instability with  $|\omega_r| \ll \gamma \ll \Omega_i$  and  $0 < k_z \ll k_y \ll a_i^{-1}$ . However, the only kinetic treatment of this mode in the literature is a brief study by Mikhailovskii and Pogutse (1966), so that the  $ka_i \gtrsim 1$  properties of this mode are not well established.

**3.  $\underline{E} \perp \underline{B}$ :** A third potential source of free energy in auroral ionospheric plasmas is an electric field perpendicular to  $\underline{B}$ . If one considers a uniform  $\underline{E}$  and a collisionless plasma, the  $\underline{E} \times \underline{B}$  drift velocity is the same for all plasma species, and no free energy is present. However, as one descends in altitude, collision frequencies increase and the combination of a perpendicular  $\underline{E}$  and collisions can lead to several different sources of free energy.

Above 150 km, where the ion-neutral collision frequency is much smaller than the ion cyclotron frequency, ion-neutral collisions can produce a highly anisotropic ion distribution when  $E \gtrsim 100 \text{ mV/m}$  (Barakat et al., 1983, and references therein). Ott and Farley (1975) pointed out that

such distributions, which become toroid-shaped for sufficiently strong electric fields, could drive a variety of ion anisotropy instabilities, and identified the Post-Rosenbluth instability with  $\Omega_i < \omega_r < \Omega_{ih}$  (lower-hybrid frequency) as the mode most likely to grow in the auroral ionosphere. St-Maurice, (1978) carried out a detailed linear stability analysis of this mode and showed, in particular, that increasing density enhances the growth of this instability. This mode is expected to be unstable in the altitude range 150-400 km, but at present there are no clear identifications of enhanced micro-fluctuations associated with this instability.

At lower altitudes (in the E region) where the ion collision frequency,  $\nu_i \gg \Omega_i$  the ion drift in the  $\underline{E} \times \underline{B}$  direction becomes much reduced and, as long as the electron collision frequency,  $\nu_e \ll |\Omega_e|$ , the electron  $\underline{E} \times \underline{B}$  drift  $\underline{v}_E$  corresponds to a relative electron-ion drift which is the potential source of a plasma instability.

Farley (1963) and Buneman (1963) used kinetic and fluid theories, respectively, to describe the instability which arises when  $v_E > c_s$  (ion acoustic speed) in this situation; thus the mode has become known as the Farley-Buneman instability. Kinetic theory is necessary to obtain the maximum growth rate of this instability (Schmidt and Gary, 1973); the linear theory in this framework has been studied by Lee et al. (1971), Ossakow et al. (1975), and Schlegel and St-Maurice (1982). The Farley-Buneman instability generally satisfies  $\gamma < \omega_r < \Omega_{ih}$  and has maximum growth rate at  $0 < k_z < k_y$ .

There is modest agreement between the radar observations of relatively short wavelength irregularities in the equatorial electrojet and the theoretical properties of the Farley-Buneman instability. There is considerably less correlation between theory and radar observations of high latitude E-region irregularities (Fejer and Kelley, 1980).

4.  $\underline{E} \perp \underline{B}$  and  $\nabla n \perp \underline{B}$ : If in addition to a cross-field  $\underline{E}$ , a density gradient perpendicular to  $\underline{B}$  is also present, the  $\underline{E} \times \underline{B}$  gradient drift instability can grow. This mode differs from the Farley-Buneman instability in two important ways: it has maximum growth rate at  $k_z = 0$  and can grow at  $v_E < c_s$ .

The kinetic theory for unmagnetized ions of the E region has been studied by Schmidt and Gary (1973). Kinetic effects may become important for this instability in the E region if  $\underline{E} \times \underline{B}$  drift-driven modes grow to large amplitudes and of themselves produce relatively steep density gradients.

Fluid calculations (Huba et al., 1983; Huba and Lee, 1983) have shown that an inhomogeneity in  $\underline{E}$  can stabilize this instability, and that short-wavelength modes are preferentially stabilized over longer-wavelength modes. This suggests that a sheared  $\underline{E}$  may stabilize kinetic modes, but that has not yet been demonstrated.

5. Other Sources of Free Energy: Instabilities due to other sources of free energy may also arise in the auroral ionosphere. Linear kinetic theories have been constructed for temperature drift instabilities (Gary and Abraham-Shrauner, 1981), a field-aligned ion flow shear instability (Gary and Schwartz, 1980) and a field-aligned electron flow shear (Keskinen and Huba, 1983). However, these modes are not directly associated with the cross-field density gradients and fields of ionospheric structures, and are less likely to play a role in structure dissipation.

## Nonlinear Effects

Linear theory provides a great deal of information about the frequencies, growth rates and propagation properties of the plasma instabilities, but does not yield any quantitative information about transport (e.g., how microinstabilities affect the large scale distributions of plasma density in the ionosphere). To understand this one must go to nonlinear theory.

Nonlinear theories for weakly turbulent instabilities may generally be classified as wave-wave (involving the exchange of momentum and energy among various plasma modes) and wave-particle (involving the exchange of momentum and energy between modes and plasma particles). Since plasma transport requires a particle interaction, we concern ourselves only with the latter process. We use the term "wave-particle" as being clearer and more appropriate for describing such transport than the adjective "anomalous."

A small amount of preliminary work has been done in this area. Huba and Ossakow (1981) considered wave-particle transport due to the universal drift in a one-dimensional diffusion equation to study diffusion on a monotonic background density gradient. They showed that, for sufficiently short density scale lengths, wave-particle diffusion due to the universal is much faster than classical diffusion. Bernhardt et al. (1982) also considered a similar model but assumed a sinusoidally varying background density. They showed that wave-particle diffusion can lead to sawtooth-shaped background density profiles. This shape corresponds to a short wavelength density power spectrum proportional to  $k^{-4}$ , which is considerably steeper than the  $k^{-2}$  spectra predicted by classical fluid theory.

Thus wave-particle diffusion due to the universal density drift instability is a likely source of the relatively steep density spectra recently observed at short wavelengths in the equatorial F region (Kelley, 1982; Kelley et al., 1982a, b; Basu et al., 1983; and Singh and Szuszczewicz, 1983). However, more nonlinear work has to be done in order to adapt these ideas to the high-latitude ionosphere.

## IV. AN EMERGING PICTURE

A general emerging picture of the birth, growth and decay of high-latitude ionospheric structures is developing. At scales greater than  $\approx 50$  km structure is believed to be primarily due to that identified with incoming particle precipitation patterns. This is especially true in the E region and lower F region where electrons with energies greater than 1 keV produce ionization with lifetimes of a few minutes. Electrons with energies of hundreds of eV produce upper F-region ionization enhancements which can have lifetimes of hours, allowing significant redistribution within the high latitude region.

This "initial" large scale structure is subject to a variety of free energy sources to drive various plasma macroinstabilities which in turn can populate the smaller scale size regimes both directly and indirectly through wave-wave interaction and cascading. Further structuring can be quenched by diffusive-like processes. Plasma microinstabilities can provide wave-particle diffusion, which in turn affects the decay of large-scale structures. These instabilities must operate on time-scales short when compared to the plasma lifetimes. Moreover, the time scales for growth and dissipation of structure may vary throughout the convecting lifetime. Therefore, the amount of structure at any given point depends on the past history of the flux tube.

In the scale size regime, greater than 100 km, high-latitude ionospheric structure has been identified principally by in situ satellite observations. Virtually no theoretical arguments exist which specifically treat the structure and scale sizes in this regime. One exception here would be the excitation of >100 km wavelength thermal gravity waves which couple to the ambient plasma and produce structure through collisional effects.

In the 1 to 100 km scale-size regime, high latitude structure has been identified using both in situ and ground-based observations. Irregularities in this scale size regime can be caused by structured particle precipitation, ion chemistry, and plasma instabilities. The mapping between magnetospheric and ionospheric structure may be important, but has not been assessed. It is known that D- and E-region auroral patches that appear in the equatorward part and morning sector of the auroral oval have irregular shapes and typical sizes from 10 to 50 km. The shape of the patches is coherent for time-scales of tens of minutes. Recently, Chiu et al. (1983) have constructed a theory of the spatial morningside auroral patch structures based on the modulation of the growth rates of electron cyclotron waves by the mirror instability in the plasma sheet.

The inverted-V structures, first described by Frank and Ackerson (1971), have long been identified with the discrete aurora forms because the precipitation occurs on field lines poleward of the trapped fluxes in the regions where the discrete aurora is commonly observed. Lin and Hoffman (1979) have made a systematic study of inverted-V events and report typical latitudinal widths of about 50 km. These dimensions are much larger than typical discrete arc widths and refer more to auroral bands. Theoretically, several mechanisms, e.g., anomalous resistivity, double layers, magnetic mirror electric fields, have been invoked to explain inverted-V's but little theoretical work has been published which addresses their spatial structure. In the auroral F region, ground-based and in situ observations have shown plasma structure in this scale size regime. These plasma structures have transverse scale sizes comparable to inverted-V events, and follow ambient auroral convection patterns. No theoretical work has been performed which can predict their scale sizes, but they have been shown to be unstable to convective plasma macroinstabilities.

Scintillation observations have verified the existence of 1 km scale size plasma structure associated with the convecting plasma enhancements. In addition, ground-based observations of east-west structure associated with these plasma enhancements are consistent with the nonlinear evolution of these convective plasma macroinstabilities. Furthermore, collisional ion cyclotron and velocity shear instabilities may directly contribute to this scale size regime.

In the scale size regime 100m to 1 km, experimental data have been acquired using in situ measurements. Particle precipitation and plasma instabilities contribute to this scale size regime. Discrete auroral arcs have dimensions of this scale size. In the auroral F region, it is known that wave-wave interaction and associated wavenumber cascade, can lead to further structuring and the appearance of smaller scale irregularities. In this scale size regime, structure may also be directly excited by the  $\mathbf{E} \times \mathbf{B}$  gradient-drift and current-convective instabilities.

In the 10m to 100m regime, experimental observations are limited to electric field and density measurements from satellites. This regime can be populated directly through the excitation of drift instabilities, and indirectly by cascade processes. In the auroral E region, reasonable agreement exists between density fluctuation levels in this scale size regime and those predicted by the nonlinear evolution of the gradient-drift instability (Ogawa et al., 1976).

Information about wavelengths less than 10m, primarily comes from radar backscatter and satellite electric field observations. In the auroral E region, there is good agreement between Farley-Buneman and  $E \times B$  instability theory and experimental observations. In the high-latitude F region, very little data in this short wavelength regime have been taken to date using backscatter radar owing to the vertical field line geometry.

Structure in the plasma density and/or the electric field may produce structure in the other plasma parameters. For example, the electron cooling rate with ions is a strong function of the plasma density. Thus at high latitudes a uniform heat input (solar EUV, for example) may result in electron temperature structure associated with structure in the plasma density. Also, structured particle precipitation fluxes will directly produce structure in the electron temperature. This in turn is manifested in a structured plasma scale height. In addition, significant frictional heating of F-region ions occurs when their motion relative to the much denser neutral atmosphere exceeds 500 m/s. This high relative motion also significantly increases the ion chemical loss rate (Schunk et al., 1976).

Sun-aligned F-region ridges of high density plasma in the polar cap have recently been discovered (Weber and Buchau, 1981). They are produced by long exposure to fluxes of soft electron precipitation, and have structure ranging from the order of 100 km to tens of meters. In the ridge where plasma densities of order  $10^6 \text{ cm}^{-3}$  exist, ion-neutral momentum transfer accelerates the F-region neutral atmosphere to near the plasma velocity in a fraction of an hour. Large scale regions of plasma structures with this density should thus produce comparably large scale structure in the velocity of the neutral atmosphere.

## V. FUTURE STUDIES

### THEORETICAL THRUSTS

Plasma instabilities have been shown to play an important role in the generation and evolution of high-latitude ionospheric structure. Because plasma macroinstabilities become highly nonlinear, further quantitative progress in this area can only be made through numerical modeling and simulation of the fundamental plasma equations. Simple analytic nonlinear models would also be desirable not only to give insight into the basic nonlinear physics, but also to validate the output from large numerical simulations. Since the drivers of these plasma macroinstabilities are density gradients, electric fields, neutral winds, and field-aligned currents, it is important to quantify the scale sizes of the drivers themselves. In addition, since the high latitude F-region ionosphere is threaded by magnetic field lines, one must investigate the consequences of coupling of the F region to the E region and to the topside ionosphere and magnetosphere. Moreover, since the convective motions can be spatially dependent, it is important to assess the influence of velocity shear on the macroinstabilities. This includes modeling shears in the nonlinear numerical simulations. The possible seeding role of neutral fluid gravity waves in producing large scale irregularities deserves further investigation.

The generation of irregularities with wavelengths of centimeters to tens of meters in the auroral F region has been given considerably less theoretical attention. Several free energy sources may drive these microinstabilities. The basic properties of linear ionospheric instabilities due to density gradients and cross-field electric fields are now reasonably well established. The role of field-aligned currents is still being assessed. The effects of sheared cross-field flows and a sheared magnetic field on microinstabilities in the ionosphere are not understood and deserve substantial attention.

An outstanding question is the saturated amplitudes and spectra of the small-scale irregularities and the role they play in influencing larger scale irregularities and structures through transport effects. Quasi-linear transport theory yields diffusion coefficients which are functions only of the macroscopic parameters and may be used in fluid computer simulations. However, this theory should be generalized to include collisions and other sources of free energy, especially electric fields perpendicular to the magnetic field. Two dimensional forms of the diffusion coefficient should be incorporated into a realistic ionospheric structure model code.

In the high-latitude E region, the strong electrojet currents drive plasma instabilities, which lead to both small- and large-scale structure. The linear theories of these instabilities can explain several features of the irregularities. Results from nonlinear studies of these instabilities (e.g., saturated amplitudes) are conflicting in several respects. Further analytical and numerical studies of these instabilities, especially the Farley-Buneman instability, are needed.

### EXPERIMENTAL EMPHASIS

The future direction of high-latitude experimental programs should benefit from lessons learned in the study of equatorial irregularities where a synergistic scientific approach (between experiment and theory) led to a successful understanding of major low-latitude geoplasma mechanisms. However, the high-latitude irregularities problem is substantially more complex and measurement requirements go beyond those applied at the equator. Thermal plasma temperatures, ion composition and charged-particle drift velocities must be measured to establish  $M_i$ ,  $T_e/T_i$ , relative drift velocities, and the cold plasma contribution to current systems. Full energy and pitch angle distributions should be determined, including the suprathermal regime ( $kT_e \approx 50$  eV). Electric and magnetic field measurements should have vector capabilities and density fluctuation measurements should include determination of phase velocities for waves in the drift, ion-acoustic, and ion-cyclotron regimes. In all cases, temporal and spatial resolution should be optimized within available subsystem support capability and the simultaneity in acquisition of all measured parameters should be required. In general, the simultaneous measurement of all known (and suspected) production, dissipation, and transport processes is a priority concern over the entire spectrum of scale sizes (hundreds of kms to fractions of a meter). It is only through such complete diagnosis that any important, unidentified physical process can be discovered.

A critical scale size presently identified by theory is the ion gyroradius ( $\approx 3m$ ). Therefore, experiment and theory should provide a complete description of the wavelength dependence of each parameter which can influence the seeding, driving, dissipation, and transport processes down at least to this scale.

Continuing to emphasize complementarity in measurements, we note that the growth rate of the  $E \times B$  gradient drift instability depends on the velocity difference between the ion and neutral gases. Thus, ion drift measurements must be complemented by simultaneous optical Fabry-Perot interferometer observations to assess the importance of this structuring process. The current-convective instability is driven by the current carried by thermal particles only. Thus, magnetometer measurements of total integrated current must be supplemented by instrumentation capable of determining that portion of the total current carried by the thermal plasma.

Another outstanding theoretical concern is the field-line coupling to other, perhaps distant, plasma processes and populations, e.g., E-region shorting and magnetospherically imposed structure. Thus, in situ measurements must span a large range of density, energy and pitch angle



regimes and must be augmented by simultaneous diagnostics, such as incoherent scatter, which can measure vertical as well as horizontal plasma structure. These diagnostics should be applied to active experiments, e.g., conductivity modification, to test specific hypotheses in controlled ways.

Finally, any viable theory must ultimately explain not only the amplitude of ionospheric structure, but its complex three-dimensional geometry. At present the only diagnostics that have some capability for specifying this geometry are incoherent scatter and spaced-receiver scintillation observations. However, even when combined, these instruments only span wavelengths from hundreds of kms to hundreds of meters. Thus, ultimately one must consider multiple in situ diagnostic systems (i.e., rocket, satellite, and/or satellite-satellite multipoint measurements) to unfold the multitude of space- and time-dependent processes active in the high-latitude ionosphere.

## VI. REFERENCES

- Aarons, J., A descriptive model of F-layer high-latitude irregularities as shown by scintillation observations, J. Geophys. Res., **78**, 7441, 1973.
- Akasofu, S. -I., Auroral arcs and auroral potential structure, in Physics of Auroral Arc Formation, S. -I. Akasofu and J. R. Kan, Editors, Geophys. Monograph 25, Am. Geophys. Un., p. 1, 1981.
- Barakat, A. R., R. W. Schunk, and J. P. St-Maurice, Monte Carlo calculations of the  $O^+$  velocity distribution in the auroral ionosphere, J. Geophys. Res., **88**, 3237, 1983.
- Basu, S., S. Basu, J. P. McClure, W. B. Hanson, and H. E. Whitney, High resolution topside in situ data of electron densities and VHF/GHz scintillations in the equatorial region, J. Geophys. Res., **88**, 403, 1983.
- Bernhardt, P. A., M. B. Pongratz, S. P. Gary, and M. F. Thomsen, Dissipation of ionospheric irregularities by wave-particle and collisional interactions, J. Geophys. Res., **87**, 2356, 1982.
- Bernhardt, P. A., M. F. Thomsen, and S. P. Gary, The nonlocal theory of periodic density drift instabilities, J. Geophys. Res., **88**, 938, 1983.
- Booker, H. G., and H. W. Wells, Scattering of radio waves by the F-region of the ionosphere, Terres. Magn., **43**, 249, 1938.
- Buchau, J., J. Aarons, J. P. Mullen, E. J. Weber, J. A. Whalen, H. E. Whitney, and E. E. Crampton, Amplitude scintillation studies in the polar region at 250 MHz, in Effects of the Ionosphere on Space and Terrestrial Systems, edited by J. M. Goodman, U. S. Government Printing Office, Washington, DC 1978.
- Buneman, O., Excitation of field aligned sound waves by electron streams, Phys. Rev. Lett., **10**, 285, 1963.
- Burke, W. J., M. S. Gussenhoven, M. C. Kelley, D. A. Hardy, and F. J. Rich, Electric and magnetic field characteristics of discrete arcs in the polar cap, J. Geophys. Res., **87**, 2431, 1982.

- Chaturvedi, P. K., Collisional ion cyclotron waves in the auroral ionosphere, J. Geophys. Res., **81**, 6169, 1976.
- Chaturvedi, P. K., and S. L. Ossakow, Nonlinear stabilization of the current convective instability in the diffuse aurora, Geophys. Res. Lett., **6**, 957, 1979.
- Chaturvedi, P. K., and S. L. Ossakow, The current convective instability as applied to the auroral ionosphere, J. Geophys. Res., **86**, 4811, 1981.
- Chaturvedi, P. K., and S. L. Ossakow, Effect of an electron beam on the current-convective instability, J. Geophys. Res., **88**, 4119, 1983.
- Chiu, Y. T., M. Schultz, J. F. Fennell, and A. M. Kishi, Mirror instability and the origin of morningside auroral structure, J. Geophys. Res., **88**, 4041, 1983.
- Clark, D. H., and W. J. Rait, The global morphology of irregularities in the topside ionosphere as measured by the total ion current probe on ESRO-4, Planet. Space Sci., **24**, 873, 1976.
- Costa, E., and M. C. Kelley, On the role of steepened structures and drift waves in equatorial spread F, J. Geophys. Res., **83**, 4359, 1978a.
- Costa, E., and M. C. Kelley, Linear theory for the collisionless drift wave instability with wavelengths near the ion gyroradius, J. Geophys. Res., **83**, 4365, 1978b.
- Curtis, S. A., W. R. Hoegy, L. H. Brace, N. C. Maynard, M. Sugiura, and J. D. Winningham, DE-2 cusp observations: role of plasma instabilities in topside ionospheric heating and density fluctuations, Geophys. Res. Lett., **9**, 997, 1982.
- Dyson, P. L., J. P. McClure, and W. B. Hanson, In-situ measurements of the spectral characteristics of F-region ionospheric irregularities, J. Geophys. Res., **79**, 1497, 1974.
- Farley, D. T., A plasma instability resulting in field-aligned irregularities in the ionosphere, J. Geophys. Res., **68**, 6083, 1963.
- Fejer, B. G., and M. C. Kelley, Ionospheric irregularities, Rev. Geophys. Space Phys., **18**, 401, 1980.
- Foster, J. C., and J. R. Burrows, Electron fluxes over the polar cap 1. Intense keV fluxes during poststorm quieting, J. Geophys. Res., **81**, 6017, 1976.
- Frank, L. A., and K. L. Ackerson, Observations of charged particle precipitation in the auroral zone, J. Geophys. Res., **76**, 3612, 1971.
- Fremouw, E. J., C. L. Rino, R. C. Livingston, and M. C. Cousins, A persistent subauroral scintillation enhancement observed in Alaska, Geophys. Res. Lett., **4**, 539, 1977.
- Gary, S. P., and B. Abraham-Shrauner, Three electrostatic temperature drift instabilities, J. Plasma Phys., **25**, 145, 1981.
- Gary, S. P., and S. J. Schwartz, An electrostatic flow shear instability, J. Geophys. Res., **85**, 2978, 1980.

- Gary, S. P., P. A. Bernhardt, and T. E. Cole, Density drift instabilities and weak collisions, J. Geophys. Res., **88**, 2103, 1983.
- Greenwald, R. A., W. Weiss, E. Nielsen, and N. P. Thompson, STARE: A new radar auroral backscatter experiment in Northern Scandinavia, Radio Sci., **13**, 1021, 1978.
- Hanuise, C., J. P. Villain, and M. Crochet, Spectral studies of F region irregularities in the auroral zone, Geophys. Res. Lett., **8**, 1083, 1981.
- Heppner, J. P., Polar cap electric field distributions related to the interplanetary magnetic field direction, J. Geophys. Res., **77**, 4877, 1972.
- Huba, J. D., and L. C. Lee, Short wavelength stabilization of the gradient drift instability due to velocity shear, Geophys. Res. Lett., **10**, 357, 1983.
- Huba, J. D., and S. L. Ossakow, On the generation of 3-m irregularities during equatorial spread F by low frequency drift waves, J. Geophys. Res., **84**, 6697, 1979.
- Huba, J. D., and S. L. Ossakow, Influence of magnetic shear on the current convective instability in the diffuse aurora, J. Geophys. Res., **85**, 6874, 1980.
- Huba, J. D., and S. L. Ossakow, Diffusion of small-scale density irregularities during equatorial spread F, J. Geophys. Res., **86**, 9107, 1981.
- Huba, J. D., and S. L. Ossakow, Physical mechanism of the lower-hybrid-drift instability in a collisional plasma, J. Atmos. Terr. Phys., **43**, 775, 1981a.
- Huba, J. D., and S. L. Ossakow, On 11-cm irregularities during equatorial spread F, J. Geophys. Res., **86**, 829, 1981b.
- Huba, J. D., P. K. Chaturvedi, S. L. Ossakow, and D. M. Towle, High frequency drift waves with wavelengths below the ion gyroradius in equatorial spread F, Geophys. Res. Lett., **5**, 695, 1978.
- Huba, J. D., S. L. Ossakow, P. Satyanarayana, and P. N. Guzdar, Linear theory of the  $E \times B$  instability with an inhomogeneous electric field, J. Geophys. Res., **88**, 425, 1983.
- Kadomtsev, B. B., Plasma Turbulence, Academic Press, New York, 1965.
- Kadomtsev, B. B., and A. V. Nedospasov, Instability of the positive column in a magnetic field and the anomalous diffusion effect, J. Nucl. Energy, Part C., **1**, 230, 1960.
- Kelley, M. C., Nonlinear saturation spectra of electric fields and density fluctuations in drift wave turbulence, Phys. Fluids, **25**, 1002, 1982.
- Kelley, M. C., and J. P. McClure, Equatorial spread F: A review of recent experimental results, J. Atmos. Terr. Phys., **43**, 427, 1981.
- Kelley, M. C., and C. W. Carlson, Observation of intense velocity shear and associated electrostatic waves near an auroral arc, J. Geophys. Res., **82**, 2343, 1977.

- Kelley, M. C., K. D. Baker, J. C. Ulwick, C. L. Rino, and M. J. Baron, Simultaneous rocket probe, scintillation, and incoherent scatter radar observations of irregularities in the auroral zone ionosphere, Radio Sci., **15**, 491, 1980.
- Kelley, M. C., R. C. Livingston, C. L. Rino, and R. T. Tsunoda, The vertical wave number spectrum of topside equatorial spread-F: estimates of backscatter levels and implications for a unified theory, J. Geophys. Res., **87**, 5217, 1982a.
- Kelley, M. C., R. Pfaff, K. D. Baker, J. C. Ulwick, R. Livingston, C. Rino, and R. Tsunoda, Simultaneous rocket probe and radar measurements of equatorial spread F - Transitional and short wavelength results, J. Geophys. Res., **87**, 1575, 1982b.
- Kelley, M. C., J. F. Vickrey, C. W. Carlson, and R. Torbert, On the origin and spatial extent of high latitude F region irregularities, J. Geophys. Res., **87**, 4469, 1982c.
- Keskinen, M. J., and J. D. Huba, Generation of lower hybrid waves by inhomogeneous electron streams, J. Geophys. Res., **88**, 3109, 1983.
- Keskinen, M. J., and S. L. Ossakow, Nonlinear evolution of plasma enhancements in the auroral ionosphere I: long wavelength irregularities, J. Geophys. Res., **87**, 144, 1982.
- Keskinen, M. J., and S. L. Ossakow, Nonlinear evolution of convecting plasma enhancements in the auroral ionosphere II: small scale irregularities, J. Geophys. Res., **88**, 474, 1983.
- Keskinen, M. J., S. L. Ossakow, and B. E. McDonald, Nonlinear evolution of diffuse auroral F region ionospheric irregularities, Geophys. Res. Lett., **7**, 573, 1980.
- Kindel, J. M., and C. F. Kennel, Topside current instabilities, J. Geophys. Res., **76**, 3055, 1971.
- Lee, K., C. F. Kennel, and J. M. Kindel, High-frequency Hall current instability, Radio Sci., **6**, 209, 1971.
- Lehnert, B., Diffusion processes in the positive column in a longitudinal magnetic field, in Proceedings of the Second Geneva Conference on the Peaceful Uses of Atomic Energy, **32**, 349, 1958.
- Lin, C. S., and R. A. Hoffman, Characteristics of the inverted-V event, J. Geophys. Res., **84**, 1514, 1979.
- Linson, L. M., and J. B. Workman, Formation of striations in ionospheric plasma clouds, J. Geophys. Res., **75**, 3211, 1970.
- Maynard, N. C., J. P. Heppner, and A. Egeland, Intense, variable electric fields at ionospheric altitudes in the high latitude regions as observed by DE-2, Geophys. Res. Lett., **9**, 981, 1982.
- McClure, J. P., and W. B. Hanson, A catalog of ionospheric F-region irregularity behavior based on OGO 6 retarding potential analyzer data, J. Geophys. Res., **78**, 7431, 1973.
- Mikhailovskii, A. B., Theory of Plasma Instabilities, VI. 2, Consultants Bureau, New York, 1974.
- Mikhailovskii, A. B., and O. P. Pogutse, Kinetic theory of the oscillations of an inhomogeneous plasma with collisions, Sov. Phys. Tech. Phys., **11**, 153, 1966.

- Mozer, F. S., ISEE-1 observations of electrostatic shocks on auroral zone field lines between 2.5 and 7 earth radii, Geophys. Res. Lett., **8**, 823, 1981.
- Ogawa, T., H. Mori, and S. Miyazaki, Rocket observations of electron density irregularities in the antarctic auroral E region, J. Geophys. Res., **81**, 4013, 1976.
- Ossakow, S. L., Ionospheric irregularities, Rev. Geophys. Space Phys., **17**, 521, 1979.
- Ossakow, S. L., Spread F theories - A review, J. Atmos. Terr. Phys., **43**, 437, 1981.
- Ossakow, S. L., and P. K. Chaturvedi, Current convective instability in the diffuse aurora, Geophys. Res. Lett., **6**, 332, 1979.
- Ossakow, S. L., K. Papadopoulos, J. Orens, and T. Coffey, Parallel propagation effects on the type 1 electrojet instability, J. Geophys. Res., **80**, 141, 1975.
- Ossakow, S. L., M. J. Keskinen, and S. T. Zalesak, Ionospheric irregularity physics modelling, AIAA-82-0147, AIAA Aerospace Sciences Meeting, Orlando, Florida, January, 1982.
- Ott, E., and D. T. Farley, Microinstabilities and the production of short-wavelength irregularities in the auroral F region, J. Geophys. Res., **89**, 4599, 1975.
- Phelps, A. D. R., and R. C. Sagalyn, Plasma density irregularities in the high-latitude topside ionosphere, J. Geophys. Res., **81**, 515, 1976.
- Rino, C. L., R. C. Livingston, and S. J. Matthews, Evidence for sheet-like auroral ionospheric irregularities, Geophys. Res. Lett., **5**, 1039, 1978.
- Rodriguez, P., and E. P. Szuszczewicz, High-latitude F-region irregularities: Intensity and scale size distributions, Memo. Rep. 5120, July, Nav. Res. Lab., Washington, D.C., 1983.
- Rogister, A., and N. D'Angelo, Type II irregularities in the equatorial electrojet, J. Geophys. Res., **75**, 3879, 1970.
- Sagalyn, R. C., M. Smiddy, and M. Ahmed, High-latitude irregularities in the topside ionosphere based on ISIS 1 thermal ion probe, J. Geophys. Res., **79**, 4252, 1974.
- Sato, T., and T. Ogawa, Self-consistent studies of two-dimensional large-scale (~100m) electrojet irregularities, J. Geophys. Res., **81**, 3248, 1976.
- Scannapieco, A. J., S. L. Ossakow, S. R. Goldman, and J. M. Pierre, Plasma cloud late time striation spectra, J. Geophys. Res., **81**, 6037, 1976.
- Schlegel, K., and J. P. St-Maurice, Note on the parallel propagation effects of unstable Farley-Buneman waves at high latitudes, Planet. Space Sci., **30**, 315, 1982.
- Schmidt, M. J., and S. P. Gary, Density gradients and the Farley-Buneman instability, J. Geophys. Res., **78**, 8261, 1973.
- Schunk, R. W., P. M. Banks, and W. J. Raitt, Effects of electric field and other processes upon the nighttime high latitude F layer, J. Geophys. Res., **81**, 3271, 1976.

- Singh, M., and E. P. Szuszcwicz, Composite equatorial spread-F wave number spectra from medium to short wavelengths, Memo. Rep. 5088, Nav. Res. Lab., Washington, D.C., June, 1983.
- Simon, A., Instability of a partially ionized plasma in crossed electric and magnetic fields, Phys. Fluids, **6**, 382, 1963.
- Sperling, J. L., and S. R. Goldman, Electron collisional effects on lower hybrid drift instabilities in the ionosphere, J. Geophys. Res., **85**, 3494, 1980.
- St-Maurice, J-P., On a mechanisms for the formation of vlf electrostatic emissions in the high latitude F-region, Planet. Space Sci., **26**, 801, 1978.
- Stern, D. P., The origin of Birkeland currents, Rev. Geophys. Space Phys., **21**, 125, 1983.
- Sudan, R. N., and M. Keskinen, Theory of strongly turbulent two dimensional convection of low pressure plasma, Phys. Rev. Lett., **38**, 966, 1977.
- Sudan, R. N., J. Akinrimisi, and D. T. Farley, Generation of small-scale irregularities in the equatorial electrojet, J. Geophys. Res., **78**, 240, 1973.
- Sugiura, M., N. C. Maynard, W. H. Farthing, J. P. Heppner, B. G. Ledley, and L. J. Cahill, Jr., Initial results on the correlation between the magnetic and electric fields observed from the DE-2 satellite in the field aligned current regions, Geophys. Res. Lett., **9**, 985, 1982.
- Szuszcwicz, E. P., J. C. Holmes, and M. Singh, The S3-4 ionospheric irregularities satellite experiment: Probe detection of multi-ion component plasmas and associated effects on instability processes, Astrophys. Space Sci., **86**, 235, 1982a.
- Szuszcwicz, E. P., J. C. Holmes, D. N. Walker, P. Rodriguez, M. Swinney, L. Kegley, and M. Singh, An atlas of F-region structures as determined by the NRL-747/S3-4 ionospheric irregularities satellite investigation, Memo. Rep. 4862, Nav. Res. Lab., Washington, D.C., July, 1982b.
- Tanskanen, P. J., D. A. Hardy, and W. J. Burke, Spectral characteristics of precipitating electrons associated with visible aurora in the premidnight oval during periods of substorm activity, J. Geophys. Res., **86**, 1379, 1981.
- Thomsen, M. F., and S. P. Gary, A nonlocal theory of an electrostatic sinusoidal density drift instability, J. Geophys. Res., **87**, 3551, 1982.
- Vickrey, J. F., and M. C. Kelley, The effects of a conducting E layer on classical F region cross-field plasma diffusion, J. Geophys. Res., **87**, 4461, 1982.
- Vickrey, J. F., C. L. Rino, and T. A. Potemra, Chatanika/TRIAD observations of unstable ionization enhancements in the auroral F-region, Geophys. Res. Lett., **7**, 789, 1980.
- Weber, E. J., and J. Buchau, Polar cap F-layer auroras, Geophys. Res. Lett., **8**, 125, 1981.
- Weinstock, J., and R. H. Williams, Nonlinear theory of microinstabilities and enhanced transport in plasmas, Phys. Fluids, **14**, 1472, 1971.

- Winningham, J. D., and W. J. Heikkila, Polar cap auroral electron fluxes observed with ISIS 1, J. Geophys. Res., **79**, 949, 1974.
- Yaeger, D. M., and L. A. Frank, Low energy electron intensities at large distances over the earth's polar cap, J. Geophys. Res., **81**, 3966, 1976.
- Zabusky, N. J., J. H. Doles III, and F. W. Perkins, Deformation and striation of plasma clouds in the ionosphere, 2, Numerical simulation of a nonlinear two-dimensional model, J. Geophys. Res., **78**, 711, 1973.
- Zalesak, S. T., S. L. Ossakow, and P. K. Chaturvedi, Nonlinear equatorial spread F: The effect of neutral winds and background Pedersen conductivity, J. Geophys. Res., **87**, 151, 1982.





# **APPENDIX A**

## **STEERING COMMITTEE AND SUBCOMITTEE MEMBERS**

### **STEERING COMMITTEE**

Dr. Dennis Papadopoulos (Chair)  
Dept. of Physics and Astronomy  
University of Maryland  
College Park, MD 20742

Dr. Jay P. Boris  
Laboratory for Computational Physics  
Naval Reseach Laboratory  
Code 4040  
Washington, DC 20375

Dr. Edward P. Szuszciewicz  
Code 4108  
Naval Research Laboratory  
4555 Overlook Ave., SW  
Washington, DC 20375

Dr. Dixon M. Butler  
Solar Terrestrial Theory Program  
Code EE-8  
NASA Headquarters  
Washington, DC 20546

Dr. Donald J. Williams  
John Hopkins University  
Applied Physics Laboratory  
Laurel, MD 20707

Dr. Mukul Kundu  
Depart. of Physics and Astronomy  
University of Maryland  
College Park, MD 20742



**SUBCOMMITTEE MEMBERS****(1) Ionospheric Physics**

Dr. Herbert C. Carlson  
Air Force Geophysics Laboratory/PHY  
Hanscom Air Force Base  
Bedford, MA 01731

Dr. Rod Heelis  
Center for Space Sciences  
University of Texas at Dallas  
P.O. Box 688  
Richardson, TX 75080

Dr. S. Ossakow (Chair)  
Plasma Physics Division  
Naval Research Laboratory  
4555 Overlook Avenue, SW  
Washington, DC 20375

Dr. Robert W. Schunk  
CASS-UMC 34  
Utah State University  
Logan, UT 84322

Dr. James F. Vickrey  
SRI International  
333 Ravenwood Avenue  
Menlo Park, CA 94025

**(2) Magnetospheric Physics**

Dr. Roger Anderson  
Department of Physics and Astronomy  
University of Iowa  
Iowa City, IA 52242

Dr. Daniel N. Baker  
Space Plasma Physics Group  
Mail Stop D438  
Los Alamos National Laboratory  
Los Alamos, NM 87545

Dr. Joe Fennell (Chair)  
The Aerospace Corporation  
M2-259  
P.O. Box 92957  
Los Angeles, CA 90009

Dr. Margaret Kivelson  
University of California, Los Angeles  
Dept. of Earth and Space Sciences  
Los Angeles, CA 90024

Dr. M. Schulz  
The Aerospace Corporation  
M2-259  
P.O. Box 92957  
Los Angeles, CA 90009

**(3) Interplanetary Physics**

Dr. Aaron Barnes  
Theoretical Studies Branch  
NASA/Ames Research Center  
Code 245-3  
Moffett Field, CA 94035

Dr. Leonard Burlaga (Chair)  
Code 692  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. William C. Feldman  
Los Alamos National Laboratory  
Mail Stop D438  
Los Alamos, NM 87535

Dr. Arthur Hundhausen  
National Center for Atmospheric Research  
HAO/SMM  
P.O. Box 3000  
Boulder, CO 80307

Dr. K. Oglivie  
NASA/Goddard Space Flight Center  
Code 692  
Greenbelt, MD 20771

**(4) Solar Physics**

Dr. Steve Kahler  
Air Force Geophysics Laboratory/PHG  
Hanscom Air Force Base  
Bedford, MA 01731

Dr. Robert Rosner  
Center for Astrophysics  
Harvard College  
60 Garden Street  
Cambridge, MA 02138

Dr. R. Ulrich  
University of California, Los Angeles  
Dept. of Astronomy  
Los Angeles, CA 90024

Dr. George L. Withbroe (Chair)  
Center for Astrophysics  
Harvard College  
60 Garden Street  
Cambridge, MA 02138

# APPENDIX B

## WORKING GROUP PARTICIPANTS AND ATTENDEES

### WORKING GROUP PARTICIPANTS

**Organizer: Dr. Dixon Butler**

**Technical Chairman: Dr. Dennis Papadopoulos**

#### GROUP ONE

B.U.O. Sonnerup,  
(Chair)  
P. J. Baum  
J. Birn  
S. W. H. Cowley  
T. G. Forbes  
A. B. Hassam  
S. W. Kahler  
W. H. Matthaeus  
W. Park  
G. Paschmann  
E. R. Priest  
C. T. Russell  
D. S. Spicer  
R. Stenzel

#### GROUP TWO

R. Rosner, (Chair)  
E. Chupp  
G. Gloeckler  
D. J. Gorney  
S. M. Krimigis  
R. Ramaty  
D. W. Swift  
L. Vlahos  
E. G. Zweibel

#### GROUP THREE

J. P. Boris, (Chair)  
C. R. DeVore  
L. Golub  
R. F. Howard  
B. C. Low  
N. R. Sheeley Jr.  
G. W. Simon  
K. C. Tsinganos

#### GROUP FOUR

A. Barnes, (Chair)  
M. Goldstein  
J. Hollweg  
J. Mariska  
W. H. Matthaeus  
E. Smith  
R. Stein  
G. Withbroe  
R. Woo

#### GROUP FIVE

J. K. Alexander,  
(Chair)  
L. F. Bargasze  
J. L. Burch  
T. E. Eastman  
J. G. Lyon  
J. D. Scudder  
T. W. Speiser  
G. -H. Voigt  
C. -C. Wu

#### GROUP SIX

A. J. Hundhausen,  
(Chair)  
L. F. Burlaga  
W. C. Feldman  
J. T. Gosling  
E. Hildner  
L. L. House  
R. A. Howard  
A. S. Krieger  
M. R. Kundu  
B. C. Low  
N. R. Sheeley  
R. S. Steinolfson  
R. T. Stewart  
S. T. Wu

#### GROUP SEVEN

Y. T. Chiu, (Chair)  
R. Anderson  
J. Fennell  
R. Hoffman  
M. Hudson  
L. Lyons  
P. Palmadesso  
E. Ungstrup  
R. Vondrak  
D. Williams  
R. Wolf

#### GROUP EIGHT

D. N. Baker, (Chair)  
S. -I. Akasofu  
W. Baumjohann  
J. W. Bieber  
D. H. Fairfield  
E. W. Hones  
B. Mauk  
R. L. McPherron  
T. E. Moore

#### GROUP NINE

L. A. Fisk, (Chair)  
R. L. Arnoldy  
L. J. Lanzerotti  
R. Lin  
E. Oran  
J. B. Reagan  
M. Schultz  
B. Tsurutani

#### GROUP TEN

E. Greenstadt,  
(Chair)  
V. Formisano  
C. Goodrich  
J. T. Gosling  
M. Lee  
M. Leroy  
M. Mellott  
K. Quest  
A. E. Robson  
P. Rodriguez  
J. Scudder  
J. Slavin  
M. Thomsen  
D. Winske  
C. S. Wu

#### GROUP ELEVEN

R. Schunk, (Chair)  
A. R. Barakat  
H. Carlson  
J. B. Evans  
J. Foster  
R. Greenwald  
T. Potemra  
M. H. Rees  
A. D. Richmond  
R. G. Roble  
J. Scudder

#### GROUP TWELVE

S. Ossakow, (Chair)  
W. Burke  
H. Carlson  
S. P. Gary  
R. Heelis  
M. Keskinen  
N. Maynard  
C. Meng  
E. Szuszczewicz, (Co-Chair)  
J. Vickrey



## ATTENDEES

Dr. Joseph K. Alexander  
Code 695  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Roger R. Anderson  
Department of Physics & Astronomy  
University of Iowa  
Iowa City, IA 52242

Dr. Roger L. Arnoldy  
University of New Hampshire  
Department of Physics  
Durham, NH 03824

Dr. Daniel N. Baker  
Space Plasma Physics Group  
Mail Stop D438  
Los Alamos National Laboratory  
Los Alamos, NM 87545

Dr. A. R. Barakat  
CASS UMC 34  
Utah State University  
Logan, UT 84332

Dr. Lee Frost Bargatze  
University of California  
Institute of Geophysics and Planetary Physics  
Los Angeles, CA 90024

Dr. Aaron Barnes  
Theoretical Studies Branch  
NASA/Ames Research Center  
Code 245-3  
Moffett Field, CA 94035

Dr. Peter J. Baum  
Institute of Geophysics and Planetary Physics  
University of California/Riverside  
Riverside, CA 92521

Dr. W. Baumjohann  
MPI für Extraterrestrische Physik  
D-8046 Garching  
FEDERAL REPUBLIC OF GERMANY

Dr. John W. Bieber  
Bartol Research Foundation  
University of Delaware  
Newark, DE 19711

Dr. Joachim Birn  
Los Alamos National Laboratory  
MS D438  
Los Alamos, NM 87545

Dr. Jay P. Boris  
Laboratory for Computational Physics  
Naval Research Laboratory  
Code 4040  
Washington, DC 20375

Dr. James L. Burch  
Southwest Research Institute  
P.O. Drawer 28510  
San Antonio, TX 78284

Dr. William J. Burke  
Air Force Geophysics Laboratory  
Hanscom AFB, MA 01731

Dr. Leonard Burlaga  
Code 692  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Dixon M. Butler  
Solar Terrestrial Theory Program  
Code EE-8  
NASA Headquarters  
Washington, DC 20546

Dr. Herbert C. Carlson  
Air Force Geophysics Laboratory/PHY  
Hanscom Air Force Base  
Bedford, MA 01731

Dr. Y. T. Chiu  
The Aerospace Corporation  
Space Physics Laboratory  
M2-256  
P.O. Box 92957  
Los Angeles, CA 90009

Dr. Edward L. Chupp  
Physics Department  
University of New Hampshire  
DeMeritt Hall  
Durham, NH 03824

Dr. S. W. H. Cowley  
Department of Physics  
Imperial College of Science and Technology  
Prince Consort Road  
London SW7 UNITED KINGDOM

Dr. C. R. DeVore  
Code 4040  
Naval Research Laboratory  
Washington, DC 20375

Dr. Timothy E. Eastman  
Department of Physics and Astronomy  
University of Iowa,  
Iowa City, IA 52242

Dr. Donald H. Fairfield  
Code 695  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. William C. Feldman  
Los Alamos National Laboratory  
Mail Stop D438  
Los Alamos, NM 87535

Dr. Joe Fennell  
The Aerospace Corporation  
M2-259  
P.O. Box 92957  
Los Angeles, CA 90009

Dr. Lennard A. Fisk  
Space Science Center  
University of New Hampshire  
DeMeritt Hall  
Durham, NH 03824

Dr. Terry G. Forbes  
Department of Applied Mathematics  
North Haugh, St. Andrews University  
SCOTLAND

Dr. Vittorio Formisano  
I.F.S.  
C.N.R.  
C.P. 27  
00444 Frascati  
ITALY

Dr. John C. Foster  
MIT Haystack Observatory  
Westford, MA 01886

Dr. S. Peter Gary  
Group ESS-8, Mail Stop D438  
Los Alamos National Laboratory  
Los Alamos, NM 87545

Dr. George Gloeckler  
Department of Physics and Astronomy  
University of Maryland  
College Park, MD 20742

Dr. Melvin Goldstein  
Code 692  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Leon Golub  
S.A.O.  
60 Garden Street  
Cambridge, MA 02138

Dr. Charles Goodrich  
University of Maryland  
Astronomy Program  
College Park, MD 20742

Dr. David Gorney  
The Aerospace Corporation  
M2/260  
P.O. Box 92957  
El Segundo, CA 90009

Dr. John Thomas Gosling  
ESS-8; Mail Stop D438  
Los Alamos National Laboratory  
Los Alamos, NM 87545

Mr. Eugene Greenstadt  
TRW Systems  
R1-1176  
1 Space Park  
Redondo Beach, CA 90278



Dr. R. Greenwald  
Johns Hopkins University  
Applied Physics Laboratory  
Johns Hopkins Road  
Laurel, MD 20707

Dr. Gerhard Haerendel  
Max-Planck-Institut fuer Extraterrestrische Physik  
8046 Garching  
FEDERAL REPUBLIC OF GERMANY

Dr. Adil Hassam  
Department of Physics and Astronomy  
University of Maryland  
College Park, MD 20742

Dr. Rod Heelis  
Center for Space Sciences  
University of Texas at Dallas  
P.O. Box 688  
Richardson, TX 75080

Dr. Ernest Hildner  
NASA/Marshall Space Flight Center  
Code ES-52  
Huntsville, AL 35812

Dr. Robert A. Hoffman  
Code 696  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Joseph V. Hollweg  
Space Science Center  
Physics Department  
DeMeritt Hall  
University of New Hampshire  
Durham, NH 03824

Dr. Edward W. Hones, Jr.  
Mail Stop D438  
Los Alamos National Laboratory  
Los Alamos, NM 87545

Dr. Lewis L. House  
National Center for Atmospheric Research  
HAO/SMM  
P.O. Box 3000  
Boulder, CO 80307

Dr. Robert F. Howard  
Mount Wilson Observatory  
813 Santa Barbara Street  
Pasadena, CA 91101

Dr. Russell Howard  
Code 4173H  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Mary K. Hudson  
Space Sciences Laboratory  
University of California  
Berkeley, CA 94720

Dr. Arthur Hundhausen  
National Center for Atmospheric Research  
HAO/SMM  
P.O. Box 3000  
Boulder, CO 80307

Dr. Devrie Intrilligator  
Carmel Research Center  
P.O. Box 1732  
Santa Monica, CA 90406

Dr. Steve Kahler  
AFGL/PHG  
Hanscom AFB, MA 01731

Dr. Michael Keskinen  
Naval Research Laboratory  
Code 4780  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Allen S. Krieger  
American Science and Engineering, Inc.  
Fort Washington  
Cambridge, MA 02139

Dr. Stamatios M. Krimigis  
Johns Hopkins University  
Applied Physics Laboratory  
Johns Hopkins Road  
Laurel, MD 20707

Dr. Mukul Kundu  
Astronomy Program  
University of Maryland  
College Park, MD 20742

Dr. Lou J. Lanzerotti  
Bell Laboratories  
600 Mountain Avenue  
Murray Hill, NJ 07974

Dr. Martin Alan Lee  
Space Science Center  
University of New Hampshire  
DeMeritt Hall  
Durham, NH 03824

Dr. Marc Leroy  
DESPA Observatoire de Meudon  
92190 Meudon  
FRANCE

Dr. Robert P. Lin  
Space Science Laboratory  
University of California  
Berkeley, CA 94720

Dr. B. C. Low  
NCAR/HAO/SMM  
P.O. Box 3000  
Boulder, CO 80307

Dr. John Lynch  
Code EE-8  
NASA Headquarters  
Washington, DC 20546

Dr. John G. Lyon  
Code 4780  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Larry R. Lyons  
Space Environmental Laboratory  
R/E/SE  
NOAA  
325 Broadway  
Boulder, CO 80303

Dr. John Mariska  
Code 4175M  
Naval Research Laboratory  
Washington, DC 20375

Dr. William H. Matthaeus  
Code 692  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Barry H. Mauk  
The Johns Hopkins University  
Applied Physics Laboratory  
Johns Hopkins Road  
Laurel, MD 20707

Dr. Nelson Maynard  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Robert McPherron  
Department of Earth and Planetary Sciences  
University of California  
Los Angeles, CA 90024

Dr. Mary Mellott  
Institute of Geophysics  
University of California  
Los Angeles, CA 90024

Dr. Ching Meng  
Applied Physics Laboratory  
Johns Hopkins Road  
Laurel, MD 20707

Dr. Abraham Moghaddan-Taaheri  
Department of Physics and Astronomy  
University of Maryland  
College Park, MD 20742

Dr. Yung MoK  
University of California/Irvine  
Department of Physics  
Irvine, CA 92717

Dr. Thomas E. Moore  
Space Science Center  
University of New Hampshire  
Durham, NH 03824

Dr. Elaine Oran  
Code 4040  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. S. Ossakow  
Plasma Physics Division  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Peter J. Palmadesso  
Code 4780  
Naval Research Laboratory  
Washington, DC 20375

Dr. Dennis Papadopoulos  
Astronomy Program  
University of Maryland  
College Park, MD 20742

Dr. Wonchull Park  
Princeton Plasma Physics Laboratory  
Princeton University  
Princeton, NJ 08544

Dr. G. Paschmann  
Max-Planck-Institut für Extraterrestrische Physik  
8046 Garching  
REPUBLIC OF WEST GERMANY

Dr. Dennis Peacock  
National Science Foundation  
Atmospheric Sciences Division  
1800 G Street, N.W.  
Washington, DC 20550

Dr. Thomas A. Potemra  
Applied Physics Laboratory  
Johns Hopkins University  
Johns Hopkins Road  
Laurel, MD 20707

Dr. Kevin B. Quest  
X-1, Mail Stop E531  
Los Alamos National Laboratory  
Los Alamos, NM 87545

Dr. R. Ramaty  
Code 665  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771  
Greenbelt, MD 20771

Dr. Joseph B. Reagan  
Lockheed Aircraft Corp.  
Lockheed Palo Alto Research Center  
Building 255, Department 52-12  
3251 Hanover Street  
Palo Alto, CA 94304

Dr. M. H. Rees  
Geophysical Institute  
University of Alaska  
Fairbanks, AK 99701

Dr. A. E. Robson  
Code 4760  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Paul Rodriguez  
Code 4108  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Robert Rosner  
Center for Astrophysics  
60 Garden Street  
Cambridge, MA 02138

Dr. E. R. Schmerling  
STARLAB/SEL  
231 Durand  
Stanford University  
Stanford, CA 94305

Dr. M. Schulz  
The Aerospace Corporation  
M2-259  
P.O. Box 92957  
Los Angeles, CA 90009

Dr. Robert W. Schunk  
CASS-UMC 34  
Utah State University  
Logan, UT 84322

Dr. Jack Scudder  
Code 692  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Neil R. Sheeley, Jr.  
Code 4172  
Naval Research Laboratory  
Washington, DC 20375

Dr. George W. Simon  
Air Force Geophysics Laboratory  
Sacramento Peak Observatory  
Sunspot, NM 88349

Dr. James A. Slavin  
Jet Propulsion Laboratory  
Mail Stop 169-506  
4800 Oak Grove Drive  
Pasadena, CA 91109

Dr. Charles Smith  
Space Science Center  
DeMeritt Hall  
University of New Hampshire  
Durham, NH 03824

Dr. Edward J. Smith  
Mail Stop 169/506  
Jet Propulsion Laboratory  
California Institute of Technology  
4800 Oak Grove Drive  
Pasadena, CA 91109

Dr. Bengt Sonnerup  
Thayer School of Engineering  
Dartmouth College  
Hanover, NH 03755

Dr. Ted W. Speiser  
Department of Astro-Geophysics  
Campus Box 391  
University of Colorado  
Boulder, CO 80309

Dr. Dan Spicer  
Code 4780  
Plasma Physics Division  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Robert Stein  
Department of Physics-Astronomy  
Michigan State University  
Department of Physics and Astronomy  
East Lansing, MI 48824

Dr. Richard Steinolfson  
Department of Physics  
University of California, Irvine  
Irvine, CA 92717

Dr. Reiner L. Stenzel  
University of California  
Department of Physics, Kn1-130E  
405 Hilgard Avenue  
Los Angeles, CA 90024

Dr. R. Stewart  
Code 4160  
Naval Research Laboratory  
Washington, DC 20375

Dr. R. G. Stone  
Code 690  
NASA/Goddard Space Flight Center  
Greenbelt, MD 20771

Dr. Daniel W. Swift  
Geophysical Institute  
University of Alaska  
Fairbanks, AK 99701

Dr. Edward P. Szuszcwicz  
Code 4108  
Naval Research Laboratory  
4555 Overlook Avenue, S.W.  
Washington, DC 20375

Dr. Michelle F. Thomsen  
ESS-8, Mail Stop D438  
Los Alamos National Laboratory  
Los Alamos, NM 87545

Dr. Kanaris Tsinganos  
Harvard Smithsonian Center for Astrophysics  
60 Garden Street  
Cambridge, MA 02138

Dr. Bruce Tsurutani  
Mail Stop 169-506  
Jet Propulsion Laboratory  
4800 Oak Grove Drive  
Pasadena, CA 91109

Dr. Eigil Ungstrup  
Danish Space Research Institute  
Lundtoftevej 2, DK-2800 Lyngby  
DENMARK

Dr. James F. Vickrey  
SRI International  
333 Ravenwood Avenue  
Menlo Park, CA 94025

Dr. Loukas Vlahos  
University of Maryland  
Astronomy Program  
College Park, MD 20742

Dr. G. H. Voigt  
Department of Space Physics and Astronomy  
Rice University  
Houston, TX 77251

Dr. Richard Vondrak  
Department 52-12, Bldg. 255  
Lockheed Palo Alto Research Laboratory  
3251 Hanover Street  
Palo Alto, CA 94304

Dr. Donald J. Williams  
Johns Hopkins University  
Applied Physics Laboratory  
Laurel, MD 20707

Dr. Dan Winske  
Department of Physics and Astronomy  
University of Maryland  
College Park, MD 20742

Dr. Michael Wiskerchen  
Code EE-8  
NASA Headquarters  
Washington, DC 20546

Dr. George L. Withbroe  
Center for Astrophysics  
60 Garden Street  
Cambridge, MA 02138

Dr. Richard Wolf  
Department of Space Physics and Astronomy  
Rice University  
Houston, TX 77251

Dr. Richard Woo  
Mail Stop 238-737  
Jet Propulsion Laboratory  
1800 Oak Grove Drive  
Pasadena, CA 91103

Dr. C. S. Wu  
University of Maryland  
IPST  
College Park, MD 20742

Dr. Cheng-Chin Wu  
Physics Department  
University of California/Los Angeles  
Los Angeles, CA 90024

Dr. Shi Tsan Wu  
University of Alabama  
Department of Physics  
Huntsville, AL 35899

Dr. Ellen G. Zweibel  
University of Colorado  
Department of Astro-Geophysics  
Boulder, CO 80309



# **APPENDIX C**

## **SUMMARY OF IONOSPHERIC PHYSICS**

HIGH-LATITUDE IONOSPHERIC PLASMA DYNAMICS;  
PHENOMENOLOGY, IRREGULARITY DISTRIBUTIONS, TRANSPORT  
AND MAGNETOSPHERIC COUPLING

Edward P. Szuszczewicz  
NASA/NSF Solar Terrestrial Physics Steering Committee  
Space Plasma Diagnostics Group  
E. O. Hulburt Center for Space Research  
Naval Research Laboratory, Washington, DC 20375

### **I INTRODUCTION**

Historically the Earth's ionosphere has been the near-space laboratory for aeronomers, ion chemists and atomic physicists; but in recent years it has attracted the plasma community with interests in the many irregular ionospheric plasma states and their associated roles in magnetospheric processes and energy balance in the overall solar-terrestrial system. As a major terrestrial sink for solar and magnetospheric events, the ionosphere has features that can be traced to such seemingly remote phenomena as solar flares, radiation-belt wave-particle interactions, and magnetospheric substorms. But the ionosphere is far from a passive element in the solar-terrestrial system, for the spatial distribution of ionospheric conductivity can affect magnetospheric plasma transport and the configuration of the magnetospheric current systems. In addition, accumulating satellite measurements show the ionosphere to be an important source of ions everywhere in the magnetosphere.

While there has been an emergence of the important role of the ionosphere in solar-terrestrial physics and the development of far-reaching implications of magnetospheric-coupling processes, a coherent assessment and global understanding have only now begun to evolve. It has become clear that high-latitude ionospheric plasma physics involves a multitude of zero-order effects including space- and time-dependent particle precipitation patterns, plasma chemistry, and charged-particle transport. These zero-order processes are interwoven with higher-order kinetic effects including instabilities, field-line coupling and wave-particle interactions. The composite perspective has three contributing and interdependent elements: (1) high-latitude ionospheric plasma structure and causal mechanisms, (2) high-latitude plasma convection and transport, and (3) magnetospheric-ionospheric coupling. These elements were treated during the June 1983 NASA/NSF Solar Terrestrial Physics Workshop within the charters of the Working Groups. A summary of many of their findings is presented below.

### **II CURRENT UNDERSTANDING**

#### **GENERAL CONSIDERATIONS**

The study of near-Earth plasma processes divides itself naturally into phenomena at low-, middle-, and high-geomagnetic latitudes where fundamentally different mechanisms can be associated with the degree of magnetic-field coupling to higher altitude magnetospheric events. At low latitudes, where the geomagnetic field tends to be horizontal, the coupling to the higher altitude magnetosphere is largely through indirect effects (wind systems, electrojet currents, etc. while at

high latitudes the more vertical magnetic field promotes strong magnetospheric coupling and direct access of energetic particles. As a result, high-latitude ionosphere plasma populations are a culmination of physical processes that include EUV and energetic-particle impact ionization, ion chemistry, horizontal and field-aligned transport, gradient- and current-driven instabilities, cold plasma acceleration processes, and a host of mechanisms yet to be uncovered in unraveling the global-scale picture. To date, many of these elements have been studied separately, with only recent attempts focusing on a fully-integrated approach. Figure 1 represents a synoptic perspective on accumulated subsets of information formatted in a way intended to demonstrate the overall interdependence. Summary results will be presented within the context of each panel in Figure 1, and subsequent sections will identify a number of outstanding issues that need to be addressed.

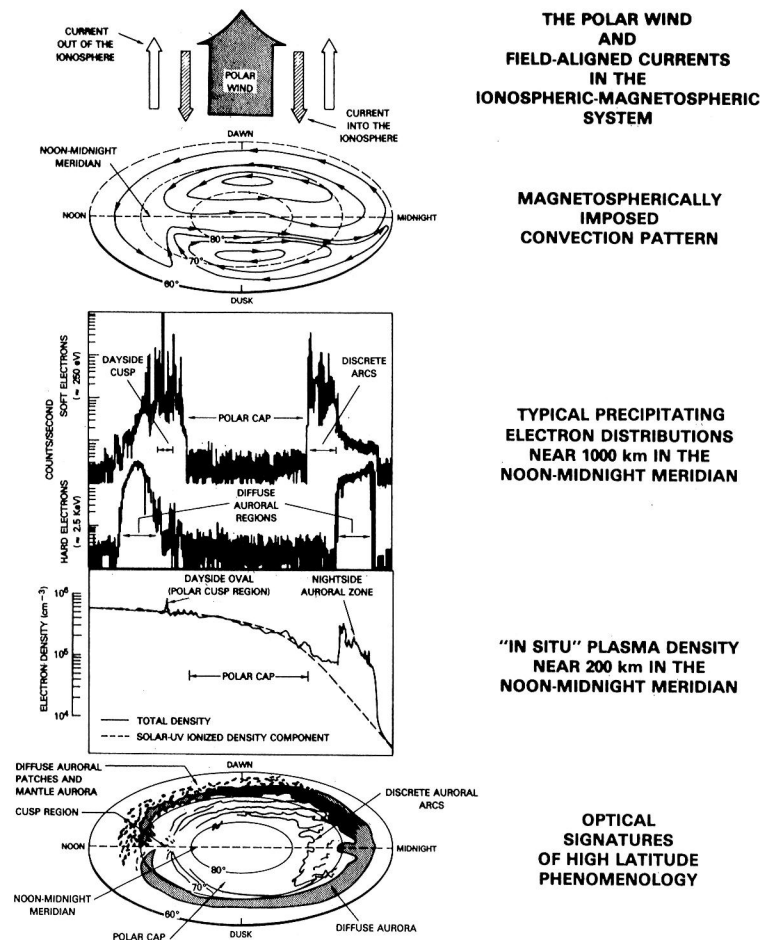


Figure 1. Composite illustration of high-latitude ionospheric processes including optically derived phenomenology, plasma density structures, precipitating particle distributions and coupling to the magnetosphere through a convection electric field, field-aligned currents and the polar wind. The format is in coordinates of magnetic local-time and magnetic latitude with specific adaptations from the works of S. -I. Akasofu (in *Physics of Auroral Arc Formation*, Akasofu and Kan, editors, Geophysics Monograph 25, Am. Geophys. U., Wash., D.C. 1982, pp. 1-14), P. Rodriguez and E. P. Szuszczewicz (*High-Latitude Ionospheric Irregularities: Intensity and Scale-Size Distributions*, NRL Memorandum Rpt. 5120, July 1983), and C. I. Meng (private communication, 1983).



## HIGH LATITUDE PHENOMENOLOGY

Ground-based and satellite-borne optical diagnostics have contributed substantially to a synoptic perspective of high-latitude phenomenology (see lowest panel, Figure 1) with major domains having been identified as diffuse auroras, discrete auroral arcs, the polar cap and the polar cusp. The statistical distribution of auroras in geomagnetic-latitude local-time coordinates has the shape of an oval belt, located near  $75^\circ$  geomagnetic latitude at noon and near  $65^\circ$  at local midnight. Diffuse aurora appear as a broad band of weak auroral luminosity with a width of several tens to hundreds of kilometers.

Discrete aurora occur along all parts of the auroral oval as well as over the polar cap. They can appear as a single faint arc or in multiple bright strands with a width ranging from a few hundred meters to a few tens of kilometers. Simultaneous measurement of precipitating electrons and optical auroral imagery show that electrons which produce discrete auroras have a mono-energetic peak as a common feature in their differential electron energy spectra. The region of diffuse aurora extends  $5^\circ$  equatorward from the region of discrete auroral arcs and forms the instantaneous auroral oval in which the discrete aurora are imbedded. In terms of magnetospheric coupling the auroral oval is thought to be the projection of the magnetotail plasma sheet into the polar ionosphere with diffuse aurora produced by direct dumping of plasma sheet electrons (see Figure 2) into the ionosphere.

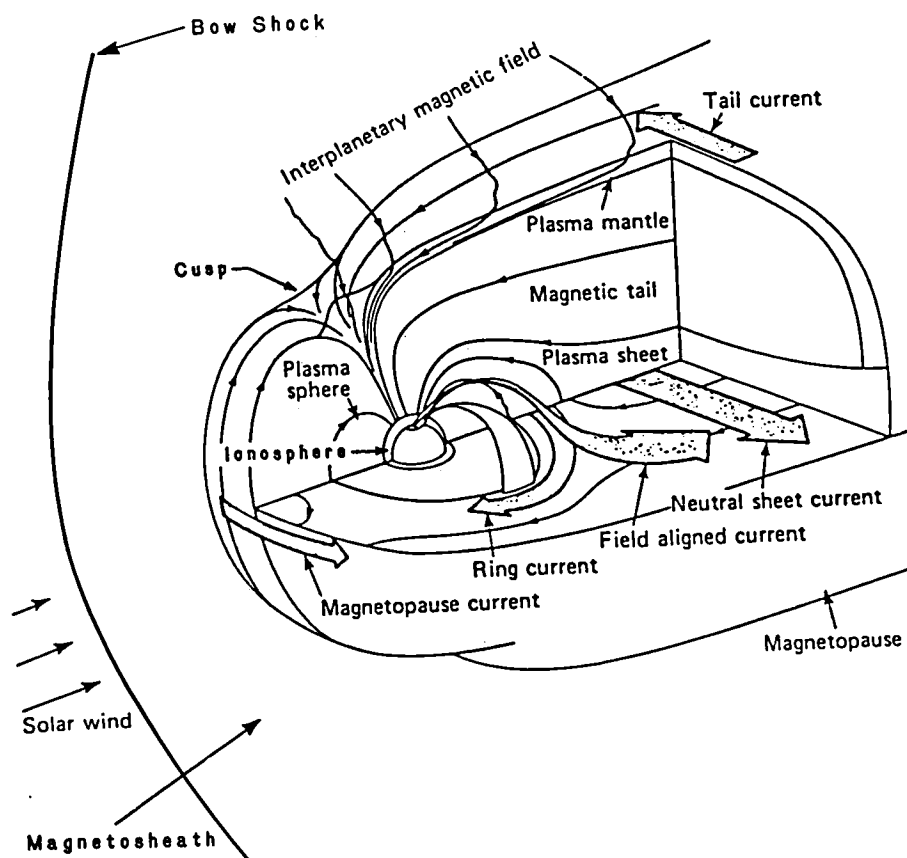


Figure 2. Configuration of the Earth's magnetosphere [adapted from "Assessment of Plasma Transport and Convection at High-Latitudes" Report of NASA/NSF Working Group #11, R. Schunk/Chairman].

The polar cusp is a limited region toward the dayside magnetosphere which extends from the high-latitude magnetopause down to the polar ionosphere. In this region, the magnetosheath plasma (Figure 2) gains direct access to the inner magnetosphere and ionosphere. The intersection of the polar cusp with the high-latitude ionosphere is statistically observed between 75° and 79° geomagnetic latitude near the noon meridian where the dayside auroral oval is situated. The particle precipitation detected in the cusp region always has a very soft differential energy spectrum.

## PARTICLE PRECIPITATION AND PLASMA STRUCTURE

Plasma in the high-latitude ionosphere is produced primarily by solar EUV ionization and precipitating electrons, with precipitating ions of substantially less consequence in terms of total plasma production. The solar EUV plasma component varies smoothly across the polar cap while the ionization component identified with energetic electrons provides the zero-order source term for plasma structure. This is illustrated in the plasma density panel in Figure 1.

Electron precipitation at high latitudes is variable in its energy distribution, horizontal extent and temporal characteristics. As a result it creates horizontal and vertical structure by virtue of its own spatial, temporal and energy configuration, and it provides a free energy source for a number of current drive and two-stream instability processes.

Electrons from the plasma sheet precipitate into the diffuse aurora domain with characteristic hard-energy spectra (a few keV) isotropically distributed in pitch angle over the downcoming hemisphere. This precipitation is usually smoothly varying (see "Precipitating Electron" panel in Figure 1) and produces diffuse auroral emission (bottom panel, Figure 1) and relatively broad, smoothly varying plasma enhancements (see "Plasma Density" panel). Variations within the diffuse auroral zone can include structured precipitation from inverted-V's and discrete arcs.

In contrast with diffuse auroral electron distributions, discrete arcs can be characterized by a distinct monoenergetic peak in the differential electron energy spectra. This spectral peak is near 1 keV for faint discrete arcs, about 3 keV for quiet discrete arcs and higher for active discrete arcs. The horizontal extent of arcs is observed from satellite-borne particle data to be always of order of 10 km to less than 100 km, while ground-based observatories report typical auroral widths of 1 km or less. The difference in observational characteristics is believed to be a beam-plasma interaction process occurring at altitudes not above several hundred kilometers.

In the polar cap, electron precipitation is referred to as "polar rain" and can generally be characterized by low energies ( $\approx 100$  eV) and low fluxes isotropically distributed outside the atmospheric loss cone. The similarity of the spectral slope of these particles to that of those measured in the dayside cusp suggests that they are of direct magnetosheath origin. The intensity of the polar rain can vary drastically but the effect on high-latitude plasma structure has generally not been of major consequence (see the polar cap region in the "Plasma Density" panel).

The particle precipitation in the cusp has a very soft differential energy spectrum with average energies near 100 eV. A characteristic feature of cusp precipitation is its highly fluctuating intensity, an observational fact not clearly identified with either a temporal nature of the incoming electrons from the magnetosheath or the filamentary structure of the precipitation. Regardless, the nature of the precipitation causes substantial "spike-like" plasma enhancements in the F region, as evidenced in the plasma density contour in Figure 1.

In general, the plasma structure produced by collective beam-ionospheric interaction processes covers a very broad range of scale sizes ranging from hundreds of kilometers to meters. Concerted efforts to characterize the spectral behavior of irregularity distributions in each of the high-latitude regions are only just beginning. A similar effort in the nighttime equatorial F region has recently contributed substantial understanding of the mechanisms responsible for the six decades of irregularity structures known to exist under conditions of spread-F.

## CONVECTIVE ELECTRIC FIELDS, CURRENTS AND LARGE SCALE PLASMA TRANSPORT

The particle-produced ionospheric plasma structure represents a primary input term for the overall process of irregularity distribution. Once created, local plasma structures are subjected to magnetic-field controls, a magnetospherically-imposed convection electric field, and interactions with superimposed current systems (see top panels of Figure 1). The electric field that convects ionospheric plasma is intimately connected with the global scale current systems flowing between the ionosphere and outer magnetosphere. The intensity and detailed structure of the electric fields and currents are determined by characteristics of the entire magnetospheric-ionospheric electrical circuit and the direction of the interplanetary magnetic field. Along with the polar wind, the convection electric field and current systems represent the fundamental link between the interplanetary medium and the near-Earth ionosphere.

All convection-electric-field measurements reported to date have provided information on only a limited spatial region at specific instants of time. As a result, the overall electric field pattern has been synthesized by combining a variety of observations in space and time. Although individual measurements indicate that the electric field pattern is structured and highly variable, there is an underlying pattern which can be approximated by two convection cells aligned near the noon-midnight meridian (Figure 1). This underlying pattern organizes charged-particle motion in the ionosphere-magnetosphere system and provides a useful basis for determining the response of the ionosphere to convection and for identifying irregularity transport properties. The general two-cell pattern of plasma convection is anti sunward in the polar cap and sunward in the morning and evening oval.

That irregularities are not always locally-produced but can be transported from other regions has been established at high latitudes by a number of observations complemented by global scale modeling. An illustration of the convection-pattern transport process is presented in the computational results of Figure 3. Those results incorporate models for convection, ion- and neutral chemistry, and local plasma production rates with a number of required input parameters (neutral composition and temperature, thermospheric winds, electron temperature, incident EUV spectra, ionization, and a magnetospheric electric field). The figure presents four altitude profiles of total plasma density in a magnetic reference frame for high-latitude winter solstice. At 160 km the combined plasma production due to solar EUV and particle precipitation is clearly evident in well-defined morphology of the isodensity contours. Since 160 km is in a molecular-ion-dominated domain ( $[\text{NO}^+]$ ,  $[\text{O}_2^+] \gg [\text{O}^+]$ ), the density contours are controlled by local production sources with transport of local plasma enhancements to other regions substantially mitigated by rapid ion-electron recombination rates. Evidence for this is seen in the polar cusp region of the plasma density panel of Figure 1 where cross polar cap convection would be expected to transport the local polar cusp enhancements far into the cap. Instead, the enhancement is locally very intense and shows evidence of substantial degradation in the cross polar cap direction. This is a typical observation at 200 km. It is also seen that there is very good agreement in the overall qualitative features of the 200 km data profile and the model results at 160 km. As altitude increases in the subsequent panels in Figure 3,  $\text{O}^+$  becomes the dominant ion and field-aligned diffusion and horizontal convection become increasingly important. Plasma structures then undergo substantial transport out of their initial production domain.

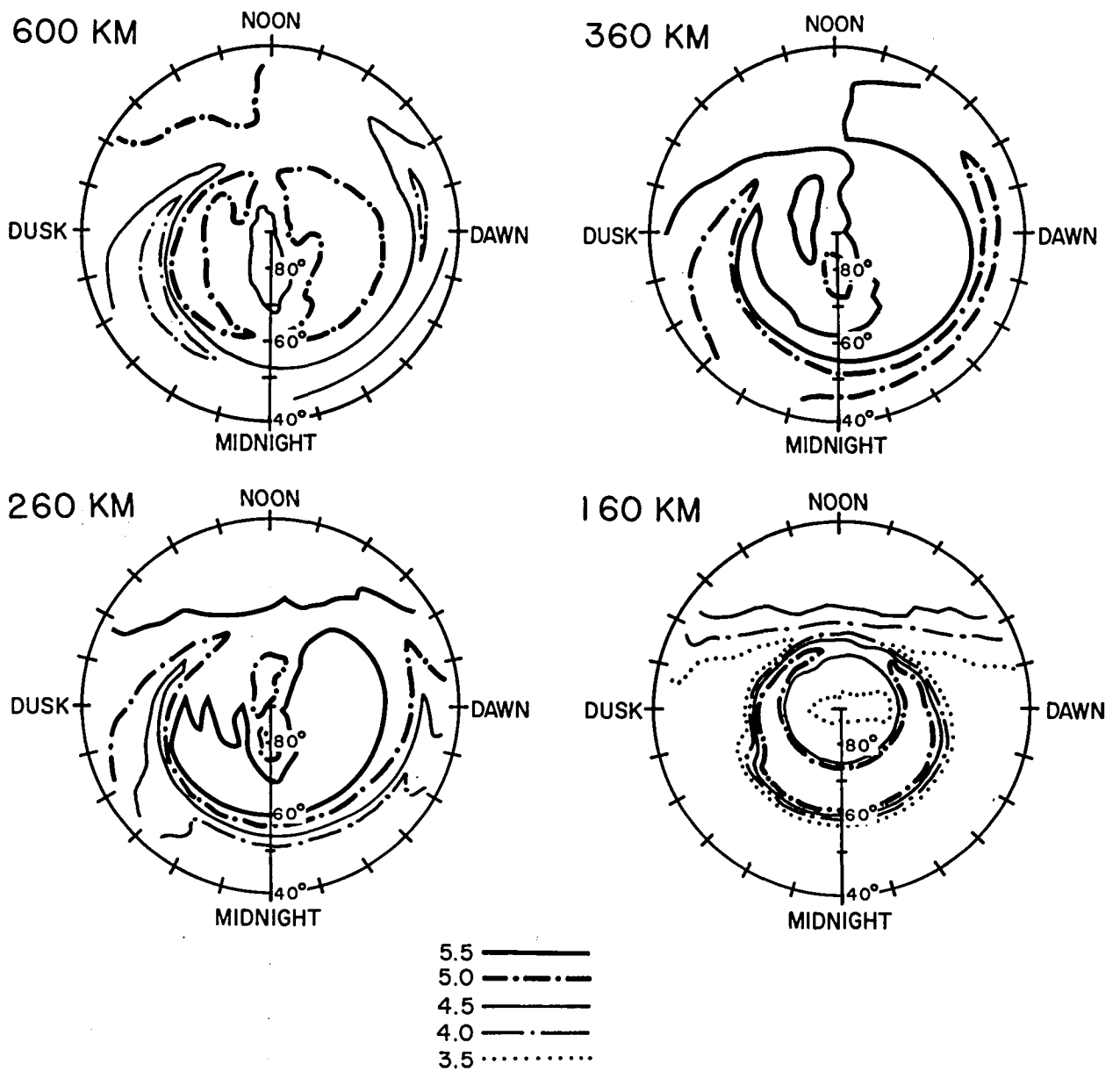


Figure 3. Ionospheric electron isodensity contours as a function of altitude in coordinates of magnetic-local-time and magnetic latitude. The legend identifies the contours in units of  $\log_{10} N_e [\text{cm}^{-3}]$ . [Adapted from Sojka, Raitt and Schunk, *J. Geophys. Res.* **86**, 6908 (1981)].

## INSTABILITY PROCESSES AND FURTHER STRUCTURING

The locally-produced and convected plasma enhancements can be subjected to a number of collective plasma instability processes either by virtue of non-equilibrium density gradients or beam-plasma, two-stream, and current-driven interactions. These instabilities can cause further structuring, affect diffusion rates, and emit a multitude of plasma waves. For high-latitude ionospheric conditions it is well known that plasma enhancements can be unstable to a number of mechanisms, including the  $\bar{E} \times \bar{B}$  gradient drift and current convective instabilities. Motivated by ground-based radar observations of convecting plasma enhancements with finger-like structures, numerical simulations have demonstrated that the  $\bar{E} \times \bar{B}$  gradient drift instability can generate plasma density and electric field structure with scale sizes ranging from tens of kilometers to tens of meters. These irregularities can take the form of striation-like structures elongated in the north-south direction for equatorward convection. Treated as an initial value problem with assumed background densities, currents, and enhancement profiles, the results of the numerical simulation are presented in Figure 4. The  $t=0$  sec panel corresponds to the assumed initial isodensity distribution in the E-W/N-S plane perpendicular to the magnetic field. The panel at  $t=1000$  sec displays the well-developed nonlinear regime where elongated medium-to-short scale structures are evident.

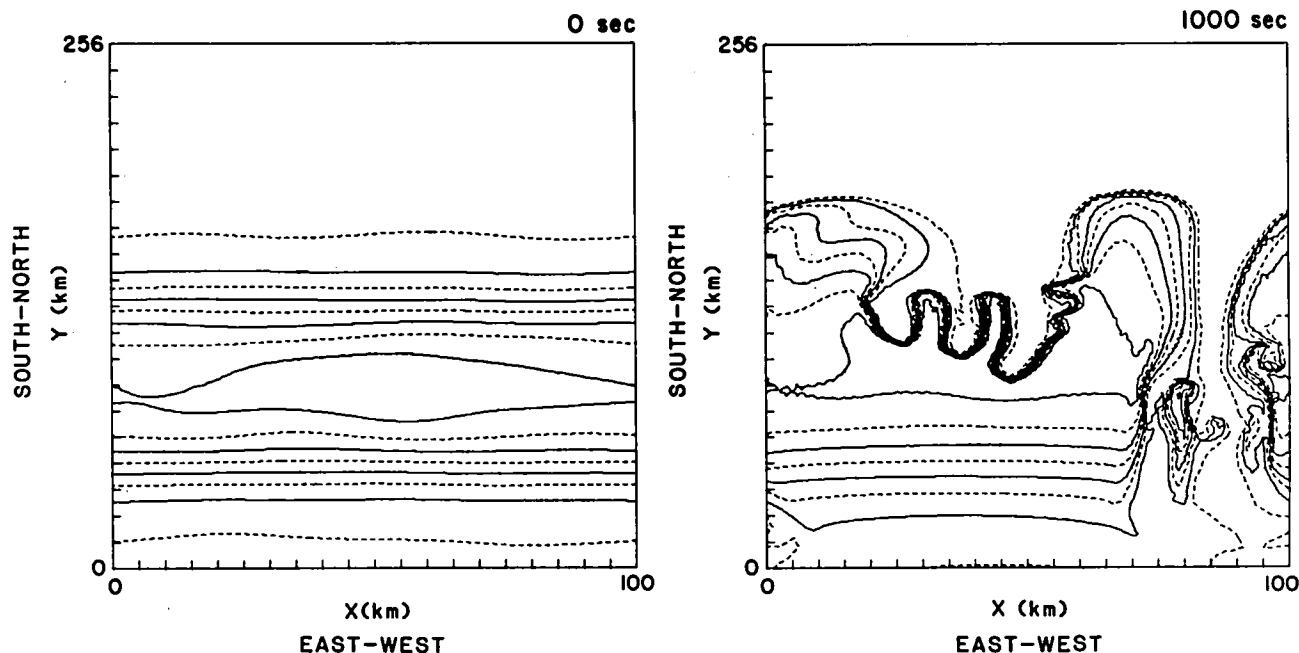


Figure 4. Numerical simulation of plasma structuring in the high-latitude ionosphere as a result of the  $\bar{E} \times \bar{B}$  gradient drift instability. Smallest grid marks correspond to 5 km spatial extents. Eight contours of dimensionless density  $n/N_0$  (where  $N_0$  zero-order background plasma density) are plotted in increments of 1.25. This gives a maximum relative plasma enhancement of 10 times that of the background ionosphere. The observer is looking upward along the magnetic field line. [Adapted from Keskinen and Ossakow, *J. Geophys. Res.* 87, 144 (1982)].

## IONOSPHERIC FEEDBACK

The feedback of high-latitude ionospheric processes onto magnetospheric phenomena can be viewed from several perspectives. The first is the role of conductivity enhancements which can modify the convection electric field, global scale current systems, and the electrodynamics of the ionospheric-magnetospheric system as a whole. Preceding sections have demonstrated the very complex nature of plasma distributions and associated conductivities, but virtually none of this information has been folded into magnetospheric models that couple to the ionosphere. Typically the ionosphere is treated as a simple linear resistor using statistically-averaged auroral conductivities. It is now clearly established that the ionosphere is a nonlinearly-active medium with conductivities highly variable in space and time and with transport characteristics that can move conductivity structures across field lines and couple to different regions of the magnetosphere.

A second feedback mechanism involves the direct transport of ionospheric ions upward along the field-lines into the magnetosphere. At high latitudes the geomagnetic field lines are not closed but extend into the tail of the magnetosphere (Figure 2). Along these open field lines light ions ( $H^+$  and  $He^+$ ) are capable of escaping from the topside ionosphere owing to the pressure difference between the ionosphere and the magnetosphere. This flow has been termed the "polar wind" and occurs over the entire polar cap (see top panel Figure 1). At present there are few observations of the thermal ions which make up the polar wind. The exact composition and its variations are not well known although it has generally been established that the polar wind is dominated by  $H^+$ , with  $He^+$  a minor constituent and with very small concentrations of  $He^{++}$ ,  $O^+$ , and  $O^{++}$ . This polar wind plasma, via cross-tail convection into the plasma sheet, provides an important source of thermal ionospheric ions in the magnetosphere.

The auroral zone can also provide hot plasma ions to the magnetosphere. This can happen through two mechanisms, both of which involve ion acceleration. One mechanism energizes ions perpendicular to  $\vec{B}$  and provides charged-particle population referred to as "ion conics". This source of hot ions has been observed to be of global scale without limitations to any local time sector of the auroral oval. The second mechanism is field-aligned acceleration of ions by electric fields  $\vec{E}$  parallel to  $\vec{B}$ , resulting in an ion beam. The combined process of transverse and field-aligned acceleration may also occur, providing an overall hotter ion distribution than in an isolated  $\vec{E}$ -parallel-to- $\vec{B}$  acceleration process.

Given that ionospheric ions are drawn out of the ionosphere and accelerated, mechanisms for trapping these particles and distributing them throughout the magnetosphere are still being sought.

## III ISSUES AND FUTURE DIRECTIONS

Despite the successful elucidation of high-latitude phenomenology and the development of perspectives on global-scale magnetospheric-ionospheric coupling, a coherent assessment of the separate components and their interdependences has yet to be made. While each of the elements has identified specific issues to be addressed (see Table 1), they all have a common thread that requires detailed understanding of spatial and temporal behaviors, plasma and energy redistribution processes including acceleration, and active wave-particle interactions. A large number of waves, including whistlers, electrostatic-ion-cyclotron, lower-hybrid and Langmuir waves have been observed and/or are theoretically predicted to occur on auroral field lines. In addition, there exists a correspondingly large variety of wave-particle interaction phenomena active in the overall system. Wave-wave and wave-particle processes are of interest in their own right as manifestations of specific phenomena, as in auroral kilometric radiation and ion conics;

they are also of interest because they can affect magnetospheric-ionospheric coupling through plasma turbulence modification of plasma transport coefficients. In addition, plasma turbulence can result in anomalous heating and subsequent production of macroscopic dynamical effects by increasing local pressure, generating potentials, etc. Clearly, wave-particle interactions and associated plasma turbulence play a critical role over a significant range of scale sizes, and the complex interactions between them require an open-minded scientific approach that carefully scrutinizes assumptions in instrument design, data analysis and theoretical modelling. As an example, the high-latitude irregularities problem is substantially complex, with measurement and computational requirements going far beyond those applied at lower latitudinal domains. Thermal plasma temperatures, ion composition and charged-particle drift velocities must be measured to establish  $M_i$ ,  $T_e/T_i$ , relative drift velocities, and the cold plasma contribution to current systems. Computational codes should carefully scrutinize assumptions of isothermal plasmas and fully-Maxwellian populations. Full energy and pitch angle distributions should be determined, including the suprathermal regime ( $kT_e \approx 50$  eV). Electric and magnetic field measurements should have vector capabilities and density fluctuation measurements should include determination of phase velocities for waves in the drift-, ion-acoustic, and ion-cyclotron regimes. In all cases, temporal and spatial resolution should be optimized within available subsystem support capabilities and the simultaneity in acquisition of all measured parameters should be optimized. In general, the simultaneous measurement of all known (and suspected) production, dissipation, and transport processes is a priority concern over the entire spectrum of scale sizes (hundreds of kms to fractions of a meter). It is only through such complete diagnoses that important, unidentified physical processes are most likely to be uncovered.

As an illustrative point of focus for future work it is easy to point to discrete auroral arcs. They represent complex phenomena entailing a spectrum of temporal and spatial scales of the  $\vec{E}$  parallel-to- $\vec{B}$  accelerating potential. As such, there is as yet no single mechanism which accounts for all observed scales. Perhaps not so surprisingly, independent attempts to find support in auroral observational data for theoretical models of field-aligned potential drops based on double-layers, electrostatic shocks, kinetic Alfvén waves, anomalous resistivity, and adiabatic processes have all had some success. As a result, there is an emerging consensus that ionospheric electrodynamics and magnetospheric coupling involve a complex interplay among several, if not all of these mechanisms. Additional studies are required which are aimed at clarifying the way in which different electrodynamic processes interact; and future efforts must be directed to identifying the regions of parameter space in which individual processes might dominate.

The emerging picture is a simple one...substantial advances have been made in studying individual component mechanisms in the high-latitude ionosphere and in its coupling to the magnetosphere; however, the high space and time variability and complex interactions among the many processes are only now beginning to be studied. It is clear that a fully satisfactory understanding of the important role of the ionosphere in the overall solar-terrestrial system can be achieved only after all the major processes have been identified and the major interactions between microscopic and macroscopic processes have been elucidated.

TABLE 1

SOME OUTSTANDING ISSUES IN HIGH-LATITUDE IONOSPHERIC PLASMA PHYSICS AND  
MAGNETOSPHERIC-IONOSPHERIC COUPLING

HIGH-LATITUDE PLASMA STRUCTURE	IONOSPHERIC PLASMA TRANSPORT AND CONVECTION	MAGNETOSPHERIC-IONOSPHERIC COUPLING
<ul style="list-style-type: none"> <li>•What are the underlying morphological distributions of plasma structures and their temporal variations?</li> <li>•How do ion composition, charged-particle temperatures, drift velocities electric fields and current systems vary along with plasma density structures?</li> <li>•What are structure scale-size distributions and associated causal mechanisms?</li> <li>•What is the role of plasma transport and what is the structure lifetime?</li> <li>•What are the roles of magnetospheric dynamics and the background ionosphere?</li> <li>•How do plasma processes couple between altitude regimes?</li> <li>•What are the spatial scale size distributions of precipitating particles and what are the charged-particle energy distribution functions from thermal to primary beam energies?</li> </ul>	<ul style="list-style-type: none"> <li>•How do potential distributions and plasma convection patterns vary spatially and temporally?</li> <li>•What is the relationship of convection patterns in conjugate ionospheres and in neutral wind systems?</li> <li>•What is the response of global ion distributions to plasma convection?</li> <li>•What is the overall polar wind contribution to magnetospheric ion populations?</li> <li>•What is the altitude distribution of polar wind ion density, composition, drift velocity, and temperature?</li> <li>•Can the atmosphere-ionosphere dynamo affect magnetospheric transport?</li> <li>•What are the precise inter-relationships among field-aligned and horizontal currents, convection electric fields, neutral winds, and ionospheric conductivities?</li> <li>•What are spatial and temporal distributions of dark-hemisphere, high-latitude ionospheric conductivities?</li> </ul>	<ul style="list-style-type: none"> <li>•What are the generators for the full spatial spectrum of field-aligned currents and what is the nature of the auroral return current?</li> <li>•How do magnetospheric-ionospheric coupling processes differ in the day and post-midnight sectors of auroral phenomenology?</li> <li>•What are the active wave-particle and wave-wave interaction in the magnetospheric-ionospheric coupling process?</li> <li>•How does ionospheric feedback influence magnetospheric phenomena?</li> <li>•What are the polar wind and ion conic acceleration mechanisms?</li> <li>•What are the distribution and trapping mechanisms for magnetospheric ions of ionospheric origin?</li> <li>•Is there a unified model of auroral potential structure and field-aligned currents which describes auroral arc formation mechanisms?</li> </ul>

### ACKNOWLEDGEMENTS

This Executive Summary has attempted to synthesize the extensive efforts of the Solar Terrestrial Physics Working Groups in the areas of "High-Latitude Ionospheric Structure" (W.G. #12), "Plasma Transport and Convection" (W.G. #11), and "Connections Between the Magnetosphere and the Ionosphere" (W.G. #7). Extensive materials have been taken directly from their individual reports in an effort to preserve the details of their perspectives on many of the contributing plasma mechanisms. Additional sources included personal notes taken during the meetings at Coolfont and from individual discussions with Drs. Ossakow, Schunk, and Chiu, the Chairmen of the three working groups. To them, and to the individual Working Group members I extend my sincere thanks for their outstanding contributions and stimulating discussions.





1. Report No. NASA RP-1120		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle SOLAR TERRESTRIAL PHYSICS: PRESENT AND FUTURE				5. Report Date June 1984	
				6. Performing Organization Code	
7. Author(s) D. M. Butler and K. Papadopoulos, Editors				8. Performing Organization Report No.	
9. Performing Organization Name and Address University of Maryland College Park, Maryland 20742				10. Work Unit No.	
				11. Contract or Grant No. NASW-3726	
				13. Type of Report and Period Covered Reference Publication	
12. Sponsoring Agency Name and Address Office of Space Science and Applications National Aeronautics and Space Administration Washington, D.C. 20546				14. Sponsoring Agency Code Code EE	
15. Supplementary Notes					
16. Abstract  This document presents a comprehensive statement of man's current understanding of solar-terrestrial physics. The discussion is organized into twelve separate topics. This document is the result of a workshop held in Berkeley Springs, West Virginia on June 6-10, 1983 at the Coolfont Recreation Center. The workshop was sponsored by the Office of Space Science and Applications, National Aeronautics and Space Administration and the National Science Foundation.					
17. Key Words (Suggested by Author(s)) Solar-Terrestrial Physics Plasma Physics			18. Distribution Statement STAR Category 92 Unclassified - Unlimited		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 590	22. Price A25		

















National Aeronautics and  
Space Administration

Washington, D.C.  
20546

Official Business

Penalty for Private Use, \$300

SPECIAL FOURTH CLASS MAIL  
BOOK

Postage and Fees Paid  
National Aeronautics and  
Space Administration  
NASA-451



**NASA**

POSTMASTER: If Undeliverable (Section 158  
Postal Manual) Do Not Return

---